CUL - H05876 - 3 - 243451

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

CALCU
CENTRAL LIBRARY
UNIVERSITY

3

FEBRUARY 1962

32
VOL. XXXII, PART I

Price Fifteen Shillings

Some Important Books

THE DEVELOPMENT OF THE INFANT AND YOUNG CHILD

By Professor R. S. Illingworth, M.D., F.R.C.P., D.P.H., D.C.H.
324 pages, 95 illustrations. 27s. 6d.

THE NATURAL DEVELOPMENT OF THE CHILD

Fourth Edition. By Agatha H. Bowley, Ph.D., F.B.Psy.S.
224 pages, 84 illustrations. 10s. 6d.

THE PSYCHOLOGICAL CARE OF THE CHILD IN HOSPITAL

By Agatha H. Bowley.
48 pages, 8 illustrations. 4s. 6d.

PSYCHOLOGY AS APPLIED TO NURSING

Second Edition. By Andrew McGhie, M.A., Ph.D.
279 pages. 17s. 6d.

E. & S. LIVINGSTONE, LTD.
TEVIOT PLACE, EDINBURGH

THE
BIRMINGHAM
PRINTERS LTD
PRINTERS & BOOKBINDERS

High Class
Technical
Publications

42-44 HILL STREET
BIRMINGHAM 5
TELEPHONE: MIDLAND 0196-7

THE EXPERIMENTAL PSYCHOLOGY OF BEAUTY

C. W. Valentine

Professor Valentine discusses the results of numerous experiments carried out on reactions to objects generally regarded as beautiful. All the visual arts, music and poetry are included in the survey.

Illustrated. 45s. net.

THE PSYCHOLOGY OF JUNG

Avis M. Dry

The author of this study is primarily concerned with Jung as a medical psychologist. She traces the development of his ideas from the experimental investigations of the 1900's to the religious and philosophic speculations of the 1950's.

35s. net.

METHUEN 36 *Essex Street* · *Strand* · *London W.C.2*

UNIVERS

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by

L. B. BIRCH

Associate Editors:

P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT
R. A. C. OLIVER
E. A. PEEL
SIR FRED J. SCHONELL
C. W. VALENTINE
S. WISEMAN

Vol. XXXII

1962

③

Published by
METHUEN AND CO., LTD., 36, Essex Street, Strand, London, W.C 2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices . 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5

CALCUTTA UNIVERSITY CENTRAL LIBRARY

EDITORIAL

The fact that this Volume begins with a new editor and a newly designed cover may alarm some readers into thinking that this betokens a change of policy for the *Journal*. They may be reassured. The pattern was set by Professor Valentine, who founded the BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY and was its editor during its first twenty-five years. That pattern was maintained by Professor Vernon who followed him, and I intend, as far as I am able, to do the same. In this I shall be greatly helped by both the former editors for as most readers know, the editorial policy of the *Journal* is guided by a group of Directors who represent the British Psychological Society and the Association of Teachers in Colleges and Departments of Education. Both Professor Valentine and Professor Vernon are Directors and, further, Professor Vernon has agreed to act as an Associate Editor and to share with me the growing burden of editorial duties.

A second Associate Editor has also been appointed in Miss Doris M. Lee, who for some years has acted as Treasurer, and now combines with it the office of Secretary, to the independent company, the British Journal of Educational Psychology, Ltd., which, in fact, owns the *Journal*. I greatly appreciate her willingness to place her scholarship and experience at my disposal.

L. B. BIRCH.

370·1505
13777

CUMULATIVE INDEX

Volumes I to XXX

A cumulative index of the articles which have appeared in the *Journal* between 1931 and 1960 is included as a supplement to this Part. It has been suggested that the supplement may be of considerable value to students and others so that a limited number of bound offprints of it will be available at 3s. each. They may be obtained from the publishers, Messrs. Methuen and Co., Ltd., 36, Essex Street, Strand, London, W.C.1. A remittance should accompany the order.

SOCIAL ATTITUDES AMONG HIGH SCHOOL STUDENTS IN INDONESIA

By BUSONO W. SUMOTIRTO

(*Department of Psychology, Gadjah Mada University, Jogjakarta, Indonesia.*)*

SUMMARY. The present investigation was carried out on the assumption that social attitudes are mainly based on the feeling of responsibility towards certain social objects. 512 subjects were taken from high school students of various Indonesian towns. They were asked to state their attitude towards fifteen situations constructed like Havighurst and Taba's Life Problem Test By means of the method of paired comparisons, the fifteen situations were linked in such a way as to get a complete set of comparisons among six kinds of responsibilities, namely towards oneself, family, friends, school, society and an abstract value (honesty).

The analysis of the students' replies was conducted according to (*a*) the direction of their choice ; (*b*) the reasons underlying their choice ; and (*c*) the difference between their replies on the normative level (' what should be done ') and the realistic level (' what would be done ').

From the results of the analysis in the main, the following conclusions were drawn :

1.—It was evident that responsibility towards society had a high percentage of choices, both on the normative and the realistic level As this tendency was found in all the six types of schools investigated from the various towns, it can be said that at present, in general, the social orientation of Indonesian students is relatively high if compared with their individualistic orientation, especially in situations where they are directly faced with the interests of the society and the country.

2.—There are four main attitude types among the subjects · (1) social type ; (2) common type ; (3) individualistic type ; (4) double-valued type. The social type which puts a stronger emphasis on social values comprises the majority of cases.

I.—INTRODUCTION.

ACCORDING to the investigations of G. W. Allport and J. M. Gillespie, American and often European students are primarily concerned with ' private values ' and are absorbed in the ' goals of private worldly success.' Allport refers to this by using the term ' privatism '.[1] Similarly, Hughes and Thompson, in the United States, found that Northern white children put a lower value-emphasis on justice and group unity than Southern-Negro children.[5] Lee and others have stressed the basic nature of individualistic values in the culture patterns of societies generally designated as ' Western.'[7]

In Indonesia, however, W. F. Wertheim states that though there was a " development towards greater individualism in the first phase of the present century," it was subsequently " being overtaken by a new movement with greater emphasis on collective action ([18], p. 165). Prof. M. Sardjito has expressed the opinion that the spirit of serving society, the spirit of ' gotong-royong ' or mutual help and friendship still exists and still forms the basis of life in the villages ([11], p. 8). Moreover, experiences during the time of the revolution, in the years following 1945, have had a deep and enduring influence in creating and developing co-operative attitudes.

---

* I am grateful to my promotor, Prof. Kurt Danziger, for his valuable advice and to Prof. Philip E. Vernon for his positive criticism in improving certain parts of this paper.

On the other hand, Dr. Ki Hadjar Dewantara has pointed out that many of the young people have drifted away from the real life of Indonesian society and its cultural traditions. At the same time, they are not yet able fully to assimilate the values of Western civilization and culture (3, p. 197).

Considerations such as these have suggested an investigation as to how far the value of 'gotong-royong' or mutual help and the value of serving the community and making sacrifices, which were shown by Indonesian students during the anti-colonial struggle, are still influencing the social attitudes of the present generation of students. The investigation was carried out to test the following hypotheses :

(1) There is a tendency among Indonesian students to stress social relative to individualistic values, especially if they are faced with hypothetical situations which directly concern the well-being of the society or country.

(2) This social attitude differs with sex, social strata, and concrete conditions in the environment.

Subjects.

Due to technical considerations, the empirical investigation was limited to high school students. In view of the specific situation in Indonesia, where socio-economic and cultural influences are different in the various regions, a regional sampling method was used in this investigation. In total, 512 high school students were included as subjects taken from various towns and schools in Indonesia :

(*a*) *Djakarta*, the capital with a metropolitan milieu ; two schools sampled. *Djakarta I* is a regular Government school, where the majority of the students comes from 'upper class' families, with an average age of 17·6, while *Djakarta II* is a private evening school with students coming from middle or lower class families : many of the students work in the morning, or do any kind of job to earn their living. Their average age is 19·4.

(*b*) *Jogjakarta*, in Central Java, with a traditional cultural milieu , two schools sampled. *Jogjakarta I* is like Djakarta I a regular Government school, with students generally coming from the 'upper class'. *Jogjakarta II*, on the other hand, is a private school with students from middle or lower class families. It has a special kind of educational system, which very much resembles that of Tagore in India, putting a strong emphasis on national culture and the social aspects of education.

(*c*) *Kediri*, in East Java, with an agrarian setting ; one school sampled. Here also the sample is a private school laying stress on social education by adjusting its curriculum to the needs of the rural environment.

(*d*) *Bandjarmasin*, in South Borneo, with a strong religious emphasis ; one school sampled. A Government school is taken here, but in fact in this 'closed community' everything is very much bound to the traditional patterns of life.

The average age of the subjects ranged from 17·6 to 19·4. Of the total of 512 subjects there were 314 boys and 198 girls. Since they were taken by random from all the three different classes in each school, it was assumed that they would give a fair reflection of the general attitude prevailing in each school.

II.—METHOD.

On the assumption that social attitudes are mainly based on the feeling of responsibility towards certain social objects, the subjects were asked to make paired comparisons between a set of six such social objects.

The comparisons were modelled on Havighurst and Taba's Life Problem Test.[4] This involved the construction of fifteen situations in such a way as to give a pair by pair comparison of six directions of responsibility, i.e., towards Self (A), Family (B), Friends (C), School (D), Society (E), and an Abstract Value like honesty (F).

Of these values, responsibility towards society represents a social attitude, while self-interest represents an individualistic one. Responsiblity towards school and honesty are also categorised as social attitudes, since both values implicitly denote a preference for social rather than individualistic norms.

The subjects were also asked ' what should be done ' (normative level) and ' what would actually be done ' (realistic level) for each situation and were requested to give their reasons for their choice each time. This method has been used by Morris[10] in his Summerfield research. Because of the large number of subjects and the large area covered, the present investigation was not based on interview material but on questionnaire responses.

The above-mentioned attitude scales centred around responses to the fifteen problem-situations were of the following type:

(1) N. has to go for rehearsal prior to facing her first piano-solo performance. This is especially urgent since during the last few days she was not able to practice. But on that same day her friends asked her to go to school to help to clean the class-room, because the Inspector was coming to see the school the next morning.

(A) *What should N do?*

Better to go for rehearsal. The cleaning of the school can be left to her school-mates.

Better to go to school and join her friends in cleaning the dirty class-rooms.

(B) *Which would N actually do?*

(2) During the final examination J. has seen his best friend K. copying. The teacher suspected K. and asked J. whether K. had copied or not. J. knew that if he told the truth, his friend K. would be expelled and would not be able to continue his studies, since K's parents were very poor. At school J. is known as Honest John; the teacher had full confidence in him and stressed to J. that he would keep the secret for himself.

(A) *What should J. do?*

Tell the teacher he did not know at all, since at that time he was himself too busy with his own examination.

Tell the teacher the truth that K. has copied during the examination.

(B) *Which would J. actually do?*

III.—RESULTS.

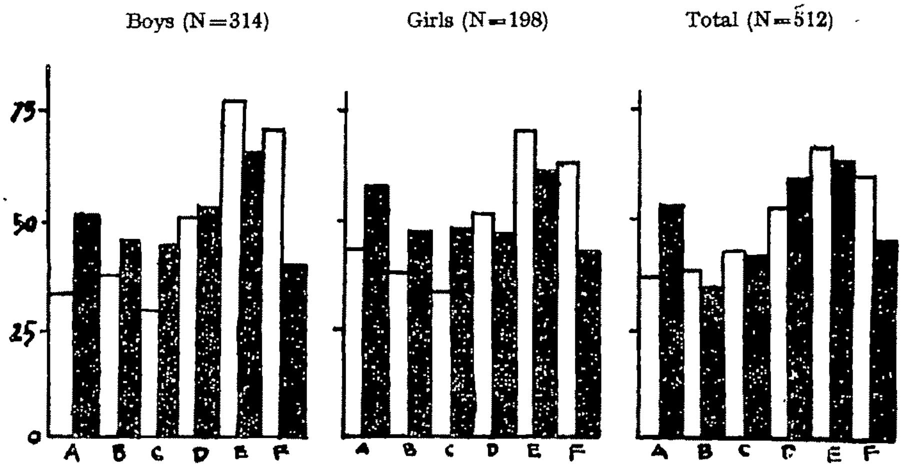
The responses of the students were analyzed (*a*) according to the direction of choice; and (*b*) according to the motivation of choice, as was done by Morris.[10] Furthermore, on the basis of the predominance of the various values on both the normative and the realistic level different attitudinal types can be distinguished.

Analysis of the direction of choice.

The comparison of the values according to the directions of choice can be shown by means of the graphic representation in Figure 1.

FIGURE 1

Percentage Choice of Each of the Six Values.



☐ : normative level. ■ : realistic level.

A=Self.
B=Family.
C=Friends.
D=School.
E=Society.
F=Abstract Value (Honesty).

Note.—50 per cent. choice means that the value was chosen in preference to some other value on half the occasions.

From these diagrams it can be seen that the score for society (E) is always very high, at the normative (at least 73 per cent.) as well as at the realistic level (at least 60 per cent.). The difference between the score for Society (E) and for self (A) is very large at the normative level (28 per cent. to 53 per cent.), while at the realistic level this difference is smaller, though still positive (5 per cent. to 35 per cent.). This fact is very important in relation with Hypothesis 1 mentioned above.

It is also interesting to note that the score for School (D) in general remains constant (around 50 per cent.), at the normative as well as at the realistic level.

The position of friends (C) which at the normative level is lower than Family (B), occupies a higher place at the realistic level. This can be understood when we realize the role of the peer-group in adolescent life. At this age the adolescent often tends to listen more to the views of his peer-group than to his parents.

There is a marked difference between the attitudes at the normative and those at the realistic level. At the normative level the scores for Society (E) and for Abstract Value (F) are high and the scores for Self (A), Family (B) and Friends (C) are low, while at the realistic level the scores for Society and Abstract value decrease, and those for Self, Family and Friends increase. These findings are in accordance with the conclusions of Morris in his Summerfield research.

Analysis of the reasons for choice.

In Hilda Taba's Life Problem Test ([4], p. 265), the subjects were given alternative answers (13-18 alternative reasons for each question), which were, in fact, based on six kinds of values, i.e., honesty, responsibility, moral courage, loyalty, friendliness and selfishness. Each alternative reason represented different values, and the subjects had to choose the reason that seemed to them to be the best. It would appear that although this technique can make quantitative analysis easier, it is not suitable for our purposes : (1) Indonesian students are not yet used to such techniques and the presence of so many different alternative reasons may confuse them, and (2) such a technique does not give freedom to the subjects to express their own opinions in their own words. For example, it excludes some commonly given affective motives.

In his Summerfield research, J. F. Morris[10] used six categories of reasons, i.e., normative, authority, reciprocity, self-interest, independence and conformity. In the present investigation, twenty-nine different kinds of reasons were distinguished in the scoring, but, in fact, most of the replies clustered around certain categories, while other categories were less popular and had a very low frequency.

In fact only six values turned out to be popular among the students, i.e. were mentioned by more than 10 per cent. of the sample in any particular school. Values with a frequency consistently less than 10 per cent. of the total of choices were disregarded. Table 1 indicates only those values having a frequency of more than 10 per cent. In order to make comparisons easy all the frequencies are given in percentages.

In order of preference, were the the following values :

(1) The most popular value is ' responsiblity towards society or country ' at the realistic and still more at the normative level.

(2) The second popular value is ' responsibility towards school,' also both at the normative and the realistic level.

(3) Value ' honesty ' is only popular at the normative level (for the boys less than for the girls), while

(4) Self-interest and

(5) ' Kearifan ' : ' prudence ' or ' enlightened-self-interest ' are popular at the realistic level.
Value No. 4 represents motives based on self-interest or selfishness, while value No. 5 is primarily directed towards public interest, but implicitly will be beneficial also for the subject concerned (prudence or enlightened self-interest), e.g., doing social work in order to become popular among his friends.

(6) Value ' tepaslira ' : ' sympathy ' or ' shared feelings with others ' has a low score and is to be found chiefly among the girls. It represents a kind of social feeling, where the subject places himself in the situation of others. The feeling usually is spontaneous in character, without any selfish background whatsoever.

A high degree of responsiblity towards society is regarded as expressing a high degree of social attitude, in contrast to self-interest, which denotes individualistic attitude. Responsibility towards school and honesty may also denote social attitudes, since these values implicitly to a large extent exclude individualistic considerations. From these answers it was evident that preference was given to social values rather than to individualistic ones.

TABLE 1

Popularity of the Values Expressed in Reasons Given by Boys (B) and Girls (G) at the Normative (N) and the Realistic (R) Level.

(Figures in the table indicate the percentage of the total number of choices made by the group which were devoted to the value given at the head of the column.)

Towns	Values		Society	School	Honesty	Self-Interest	Prudence	Sympathy
I.	B N=72	N	17·4	11·7	11·9	v	v	v
		R	14·5	v	v	13·6	11·6	—
	G N=50	N	17·3	15·5	11·9	v	11·7	v
		R	11·5	11·7	v	v	13·5	v
II.	B N=57	N	14·6	12·4	v	v	10·2	v
		R	11·2	10·0	—	14·0	12·5	v
	G N=9	N	18·5	14·8	14·1	v	v	11·1
		R	10·4	10·4	v	14·1	v	12·6
III.	B N=32	N	12·9	10·4	10·0	13·5	v	v
		R	11·7	v	—	19·4	v	v
	G N=22	N	16·8	11·8	13·7	v	v	v
		R	11·1	v	—	13·9	v	v
IV.	B N=69	N	15·5	13·1	10·7	v	11·1	—
		R	17·2	12·9	v	v	14·8	—
	G N=51	N	17·7	13·2	v	—	13·9	—
		R	18·0	11·6	v	v	15·7	—
V.	B N=48	N	16·5	14·7	15·1	—	11·9	—
		R	14·3	11·9	v	12·6	11·7	—
	G N=44	N	16·5	14·1	13·0	v	10·5	—
		R	15·2	10·6	v	11·5	11·7	—
VI.	B N=36	N	14·6	19·9	10·6	10·9	v	—
		R	14·1	12·2	v	13·9	v	v
	G N=22	N	13·0	v	12·7	v	v	10·7
		R	14·2	12·7	v	12·4	v	v
TOTAL	B N=314	N	15·3	13·7	11·2	v	v	—
		R	13·8	10·5	v	13·4	10·7	—
	G N=198	N	16·7	13·1	12·4	v	v	v
		R	13·4	10·8	v	11·0	10·7	v
	B+G N=512	N	15·9	13·4	11·8	v	v	—
		R	13·6	10·6	v	12·2	10·7	—

The figures denote percentages of choice :
v=mentioned by 5 per cent.—10 percent.
—=mentioned by less than 5 per cent.

Subjects :
B=Boys.
G=Girls.

Towns and Schools :
I=Djkarta I.
II=Djakarta II.
III=Jogjakarta I.
IV=Jogjakarta II.
V=Kediri.
VI=Bandjarmasin.

Attitude types.

On the basis of this distinction between social and individualistic values expressed on either the normative or the realistic level there are four possible attitudinal types. They are as follows:

(*a*) One group stresses social values, both at the normative and at the realistic level. This may be called the *Social-Type.*

(*b*) A second group stresses social values at the normative level, while expressing individualistic values at the realistic level. This is one of the most frequent and may be called the *Common-Type.*

(*c*) A further group always stresses individualistic values, both at the normative and the realistic level. This may be called the *Individualist Type.*

(*d*) Another group stresses individualistic values at the normative level, but social values at the realistic level. Such a type can be called the *Double-Valued Type.* At first glance this type is somewhat peculiar in that it recognizes individualistic norms, but at the level of 'what would actually be done,' it submits itself to the traditional Indonesian social norms. Perhaps this was done because of purposely wanting to show a ' pious ' attitude to one's surroundings. A comment by Kurt Lewin may apply here: " If the individual should try to diverge ' too much ' from group standards he will find himself in increasing difficulties. He will be ridiculed, treated severely, and finally ousted from the group. Most individuals, therefore, stay pretty close to the standards of the groups they belong to or wish to belong to ([8], p. 226).

On the basis of the predominant values expressed in his replies, each subject was assigned to one or other of these four types.* The distribution of these types is shown in Figure 2.

It is interesting to note that type I (the social type) comprises the majority of cases in those private schools having a special system of education with a community-oriented curriculum, while type II (the common type) was to be found in the general Government schools, especially in the large cities. It is worth mentioning here also that type III (the individualistic type) was usually found in the same milieu as type I (the social type). This suggests a certain rigidity of the norms, which drives people to one extreme or the other. It is also interesting to note that Jogja I (representing the ' upper class ') is closer in spirit to Djakarta (the capital) than to Jogja II (representing the ' middle and lower class ') with its predominance of social values. Type IV (the double-valued type) formed a small minority, especially in those places where the common type comprised the majority of cases. Although the Double-Valued type comprises only a minority, about 12 per cent. to 18 per cent. of students in the different cities, its existence is worth mentioning.

IV.—CONCLUSIONS.

From the results of the analysis, the following conclusions may be drawn:

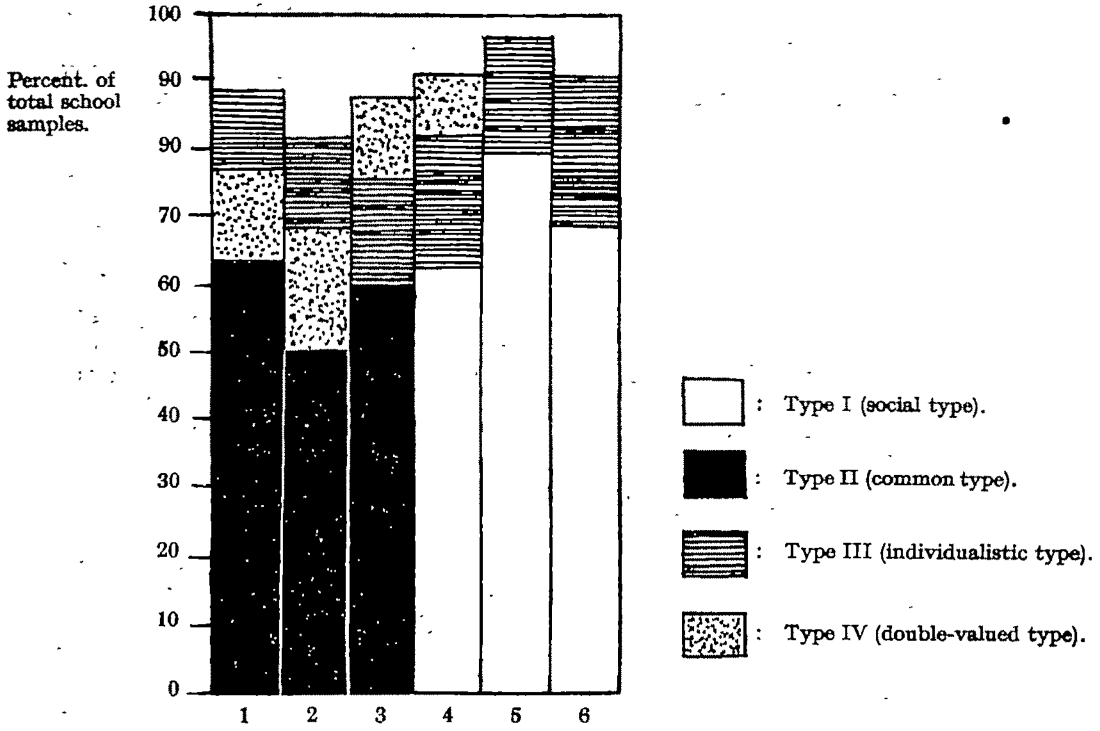
(1) From the students' replies it was evident that responsibility towards society or country had a high percentage of choices if compared with other values, both on the normative and the realistic level.

---

* Since the choice of Society itself was so popular, the assignation to the various types was primarily based on the predominance of School and Honesty on the one hand, and Self and Family on the other.

FIGURE 2

DISTRIBUTION OF THE VARIOUS ATTITUDINAL TYPES FROM THE DIFFERENT CITIES (BOYS ONLY).



1.—Djakarta I: 63 per cent. belong to Type II; 13 per cent to Type IV; 12 per cent to Type III; the rest (12 per cent.) are unspecified.†

2.—Djakarta II: 50 per cent. beong to Type II; 18 per cent. to Type IV; 13 per cent. to Type III; the rest (19 per cent.) are unspecified.

3.—Jogjakarta I: 59 per cent. Type II; 16 per cent. Type III; 12 per cent. Type IV; the rest (13 per cent.) unspecified.

4.—Jogjakarta II: 62 per cent. Type I; 19 per cent. Type III; 9 per cent. Type IV; the rest (10 per cent) are unspecified.

5.—Kediri: 79 per cent. Type I; 17 per cent. Type III; the rest (4 per cent.) are unspecified.

6.—Bandjarmasin: 68 per cent. Type I; 22 per cent. Type III; the rest (10 per cent.) are unspecified.

† " Unspecified " are those attitudes which are not clear and which cannot be classified into one of the four above-mentioned attitude types.

It is true that on the realistic level the scores for the social values decrease but they are high enough compared with the other values. As this tendency was found in all the six types of schools investigated from the various towns, it can be said that the conclusion is in line with the first hypothesis, that at present, in general, the verbalized social orientation of Indonesian students is relatively high if compared with their individualistic attitudes, expecially in situations where they are directly faced with matters involving the society and the country.

(2) A great difference can be seen if we compare the attitudes on the normative and the realistic levels. Generally, the normative values of social responsibility (E) and Abstract value (F) have a relatively high score on the normative level and decline somewhat on the realistic level; while on the contrary, the 'privatistic' values, viz.: Self-interest (A), Family (B), and Friends (C) have a low score on the normative level and a higher score on the realistic level. This conclusion is in accordance with that of the Summerfield Research of J. F. Morris.

(3) By classifying each subject according to the predominance of social or individualistic values on both the normative and the realistic level of response we can distinguish different attitudinal types: (1) social type; (2) common type; (3) individualistic type; and (4) double-valued type; the social type comprises the majority of cases. It should be pointed out also that this classification is not static or unchangeable in character, rather it is dynamic in nature.

Usually, with an attitude scale of this type, we can only study the "elicited verbal attitude"* of the subjects. Other investigations are now being undertaken to study the relationship between these attitudes and daily behaviour ('action attitude').

However, from the results mentioned above, it is evident that at the moment the Indonesian students put a strong emphasis on social values. If further empirical research should verify these findings, we would be able to confirm that the attitudes of Indonesian high school students show different tendencies to those obtaining among college students in the 'West.'

* Terms used by Green (9, p. 340).

V.—REFERENCES.

[1] Allport, G. W., and Gillespie, J. M. (1955). *Youth's Outlook on the Future.* New York: Doubleday.

[2] Dewantara, Dr. Ki Hadjar (1955). Keruntuhan dan Pem bangunan Djiwa Pemuda. Jogjakarta. *Pusara*, 11, Part XVI.

[3] Guilford, J. P. (1954). *Psychometric Methods.* New York: McGraw Hill.

[4] Havighurst, R. J., and Taba, H. (1949). *Adolescent Character and Personality.* New York: Wiley.

[5] Hughes, J. H., and Thompson, G. G. (1954). A comparison of the value systems of Southern negro and Northern white youth. *J. educ. Psychol.*, 45, 300-309.

[6] Krech, D., and Crutchfield, R. S. (1948). *Theory and Problems of Social Psychology.* New York: McGraw Hill.

[7] Lee, D. (1948). Are basic needs ultimate *J. Abnorm. Soc. Psychol.*, 43, 391-395.

[8] Lewin, K. (1951). *Field Theory in Social Science.* New York: Harper.

[9] Lindzey, G. (Ed.) (1954). *Handbook of Social Psychology* Vol. I. Boston: Addison-Wesley.

[10] Morris, J. F. (1958). The development of adolescent value judgments. *Brit. J. educ. Psychol.*, 28, 1-14.

[11] Sardjito, Prof. (1956). *Bangsa Indonesia dikemudian hari seharusnja mendjadi bangsa jang besar.* University of Gadhah Mada, Jogjakarta.

[12] Sumotirto, Busono W. (1959). *Investigation on Some Aspects of Social Attitude Among High School Students in Several Cities of Indonesia,* Ph.D. thesis: University of Gadjah Mada, Jogjakarta.

[13] Wertheim, W. F. (1956). *Indonesian Society in Transition.* The Hague—Bandung: van Hoeve.

(*Manuscript received, October 30th,* 1960.)

READING SKILL AT EIGHT YEARS IN RELATION TO SPEECH AND OTHER FACTORS

A longitudinal Study

By OLIVE C. SAMPSON
(Department of Education, University of Manchester)

Summary. This paper continues the longitudinal study of speech and language development in fifty children begun in 1956 and continued in 1959. The reading skill of these children is here the chief focus and this has been related to their earlier and present speech development. Other factors to which research has ascribed importance for their influence on reading have also been considered. The results showed an association between skill in the language aspects of speech and reading, but the contribution of the mechanical aspects of speech was unclear. Home influences, partly mediated through speech, were also found to contribute to skill and disability in reading, but personal maladjustment, though sometimes associated with reading failure and 'underfunctioning,' was not incompatible with success.

I.—Introduction.

The following study is a continuation of the work reported in 1956 and 1959. In the 1956 paper an investigation into the speech development between 18 and 30 months of fifty children representative of all levels in the community, was described. In 1959, the same fifty children were again reported on at the age of 5. Their advancement in speech skills was assessed by means of vocabulary and language tests. The results showed a positive but low correlation with the earlier speech rating. Progress seemed to be connected with the family occupational status and the children's own intelligence, as measured by the Stanford Binet Scale.

In the present follow-up focus of attention has now switched to reading. First, it is attempted to relate the reading performance of the group to their present and earlier speech development. Other factors influencing their developing skill in reading are also considered. Finally, cases of 'underfunctioning' are scrutinised with a view to throwing light on the genesis of reading retardation.

II.—Method.

All the children still available (forty-eight) were again seen. (The two lost were both boys whose earlier performances were in the better and lower average range respectively.)

Except in the case of a few private school pupils and some others who had to be seen in holiday time, the children were interviewed in school. The interviewing was done by the present writer throughout, as before. Each child was seen as near his eighth birthday as possible. Forty-three were seen within a month of that date, the remaining five within two months of it.

At the interview scores were obtained on the following tests :

(1) Neale Analysis of Reading Ability, Comprehension and Accuracy, Form C. (Neale, 1958).

(2) Mill Hill Vocabulary Test (oral forms). (Raven, 1958).

(3) Progressive Matrices (1947). (Raven, 1947).

Note was also made of the child's manner, his standard and style of speech, and his laterality.

In forty-three cases teachers completed the Bristol Social Adjustment Guide (The Child in School) (Stott, 1958).

Home visits were independently paid by a social worker (Miss Janet Craig) who made a general assessment of the home conditions, with special attention to the apparent incentives and stimulus towards school progress which the home provided. Both parents were usually seen on these visits.

III.—READING SKILL AND SPEECH.

It is often assumed, almost as self-evident, that there is a basic association between speech and reading skill. Inglis (1948) writes that " speech is one of the pre-requisites of reading." McLaren (1950) states that " in the learning of reading, speech is an important factor and the models set by parents and companions are important in reading development." Most writers who have concerned themselves with the subject have approached the problem from the point of view of the backward reader, and as Schonell (1942) points out, " the scale of speech attainment is paralleled by a scale of hearing ability," which also enters in. Some psychologists, however, appear to stress language rather than speech development, though they do not always clearly distinguish these. Betts (1954), for instance, claims that " reading functions as a facet of language development " but refers to " listening and speaking " as essential " facets of language " as a basis for reading. Gates (1954) makes the interesting suggestion that learning to understand speech and learning to read are psychologically similar tasks. But others emphasise the differences. " Nature," says Inglis (1948), " has not endowed mankind with a facility for dealing with print comparable to the facility for dealing with spoken words." M. D. Vernon (1958) clarifies the situation by separating the ' mechanical ' from the ' language ' aspect. " The relation," she says, " between reading and the mechanism of speech is uncertain," but she considers that " it is generally agreed that a child requires a good vocabulary and a competent use of language before it can learn to read."

It was the main object of this study to provide some data as to the relation between speech, language and reading. For this purpose some definitions and explanations are needed. While ' reading ' was here primarily regarded as the art of getting meaning from the printed page, the technical or barking-at-print side could not be ignored in such young readers. For this reason, Neale's Analysis of Reading Ability was used because it tests both accuracy and comprehension in a natural manner. Speech in its language aspect had been assessed at 2½ in terms of the amount of correctly phrased speech that was present (Sampson, 1956) ; at 5 it was measured by vocabulary and language tests (Sampson, 1959) ; at 8, vocabulary tests, orally administered, were again used. Cases of speech defect (using the term .loosely) noted by teachers were individually studied in relation to the reading record.

The mean scores on the reading tests proved to be closely in accordance with expectation (Neale, 1958), thus confirming the typicality of the group.These and other test results, are shown on Table 1. Owing to their date of birth, twenty-one of the children were in their third term as ' juniors,' while twenty-seven were in their first term. The scores were, therefore, studied in case this should have affected the ranking. In fact, it did not, first-term pupils obtaining a mean score of 11·5, while third-term pupils had a mean of 11·7.

TABLE 1

Test Results at C.A.8.

Test	Obtained Mean Score	Expected Score at Age
Reading Comprehension . ..	11·6	11
Reading Accuracy	31·0	28
Vocabulary............ .. .	17·6	17
Non-Verbal Intelligence	21·3	21

The meaning of the comprehension scores was further explored by correlating them with the reading accuracy score, with the vocabulary test, and with the non-verbal intelligence test. (Table 3). The comprehension results were also correlated with the speech rating at 2½ and with the vocabulary and language tests at 5 (Table 4). The correlation between the non-verbal intelligence test at 8 and the Stanford Binet (Form L) at 5 was +·57. All correlations are at least five times their probable errors.

TABLE 2

Correlation Between Reading Comprehension and Other Tests at 8 Years.

	Reading Comp.	Reading Acc.	Vocab.	Non-Verbal Intelligence
Reading—Comprehension ..	—	+·91	+·70	+·45
Reading—Accuracy ...	+·91	—	+·58	+·43
Vocabulary	+·70	+·58	—	+·60
Non-Verbal Intelligence ..	+·45	+·43	+·60	—

TABLE 3

Correlation of Reading Comprehension at 8 Years and Earlier Speech Results.

Speech at 2½	Vocabulary at 5	Language at 5
+·69	+·67	+·58

These results support the 'agreed view' that there is a considerable connection between the development of vocabulary and language and later reading skill. The high correlation between reading comprehension and accuracy was to be expected. The higher correlation between reading comprehension and vocabulary at 8 years compared with that obtained with the non-verbal intelligence test results again supported the crucial role of language in reading. On the other hand, the non-verbal intelligence test correlates quite highly with vocabulary

Following M. D. Vernon's (1958) division of the subject, it now falls to study the contribution of skill in the mechanism of speech to reading attainments in these forty-eight cases. As no standardised tests of speech exist, attention had to be concentrated on those who showed defects. Though note of speech behaviour was made at the interviews, the class teacher's rating of the child's speech was preferred. This was obtained as one item of the Bristol Social Adjustment Scale where the teacher indicated the faults of the child's speech (if any) by underlining the appropriate description. Ratings were completed for forty-three children (twenty boys, twenty-three girls). Nothing abnormal was reported in thirty-three cases (fifteen boys, eighteen girls). Of the rest, three boys and two girls were said to stutter or speak haltingly, two boys' and two girls' speech was 'thick, mumbling, or inaudible,' and one girl was 'babyish.' The sex incidence may seem surprising in view of Morley's (1957) finding of rather defective speech in 11 per cent. boys and 6 per cent. girls, respectively, but the standards on which the teacher's judgments are based probably accounts for this. As to reading, in all but two of the ten cases, the children whose speech was adversely remarked on were defective both in comprehension and accuracy. But the speech factor was not operating in isolation. For five of the eight homes involved were considerably or much below average on several counts, and in five cases a significant amount of maladjustment was present in the children. Low intelligence was a factor in three instances. Of the two whose reading was good, in spite of imperfect speech, one was a very clever asthmatic girl from a good home, the other was (according to the Bristol Scale) a rather maladjusted boy whose early speech development was markedly slow.

From this data the association between speech defect and reading skill would seem far from simple. Though it may be, as Monroe (1948) says, that " inaccurate articulation and reading disability come from a common cause " in individuals, there was certainly no single cause applicable to all these cases. The complexity of the situation became even more obvious when some of the main further factors influencing the children's reading progress were considered.

IV.—FURTHER FACTORS AFFECTING READING.

(i) *The Home Influence.*

An association has many times been found between the socio-economic status of the home and the educational progress of the child and has recently been confirmed in regard to reading by the large scale researches of Morris (1959). But the way in which this influence operates is less certain. Inglis (1948), for example, writes: " It is probably true that where socio-economic status affects reading adversely, the influence is exerted through restricted experience and through impoverished speech." But McLaren (1950), while agreeing that socio-economic status is not directly causal, puts the emphasis on " the level of mental age and of intelligence quotient and the cultural speech background." The correlation which she quotes ($\rho + 0{\cdot}31$ to $\rho + 0{\cdot}24$) is, however, low but the mention by both these writers of the speech factor is of interest in the present context.

The homes of the forty-eight children involved in this study, were more intimately known and over a longer period than is usual in researches. At an early stage (1956) the occupational status of the home and the 'stimulus to speech' which it afforded, had been noted. In the present follow-up, as the outcome of the social worker's visits, a mass of very interesting descriptive material was collected concerning the parents' attitude towards the child's school and education in general, their own reading habits, the atmosphere of the home, and the interests it provided. This was summarised in a five-point rating, independently arrived at, which indicated the amount of incentive to academic progress experienced by the child. The five-point rating and the rating by occupational status were correlated with the reading comprehension results, giving $\rho + \cdot 462(\pm \cdot 077)$ for the latter and $\rho + \cdot 497\ (\pm \cdot 067)$ for the former. Both variables are almost certainly closely related to parental intelligence and hence, probably, their similarity and importance in relation to the children's reading skill.

(ii) *Personal Adjustment.*

In Chapter Six of her book on *Backwardness in Reading*, M. D. Vernon (1958) considers the evidence that a child's adjustment affects its progress in reading. She finds this inconclusive and pleads for controlled studies and investigations " of possible temperamental or neurotic defect in cases of severe reading disability." To the writer it also seems desirable to study the incidence and history of ability and disability in relation to adjustment. The present sample gave opportunity for a contribution in this direction.

(*a*) Two estimates of the children's adjustment were obtained. (1) The teacher's estimate of the child's adjustment in school, on the Bristol Social Adjustment Scale, (2) the clinical impression obtained from the individual interviews with the psychologist.

According to the Bristol Scale, on which forty-three children were rated, all but eight (six boys and two girls) were within normal limits as defined by the group's results. Five of the eight were ' unsettled ' (or slightly maladjusted), two showed definite symptoms and one gave quite severe evidence of social maladjustment.

Turning to the clinical impressions, it was found that eight boys and six girls out of the forty-eight showed some symptoms, slight or more severe, of maladjustment of very various sorts. Six of these cases were also cases of maladjustment in school, and in three further cases the school adjustment was unknown. Four had been independently referred for Child Guidance and parents were considering referral in a fifth case.

An interesting tendency emerged when the σ scores on the reading comprehension test obtained by those deemed maladjusted either at school, at interview or otherwise, is shown on Table 4.

TABLE 4

CHILDREN DEEMED MALADJUSTED.

Reading Score	On Bristol Scale		On Clinical Evidence	Independent Child Guidance Referral
	Slight	Definite		
$+1\sigma$ N=7	2	—	5	3
Average N=3	1	1	—	—
-1σ N=6	2	2	1	1

(*b*) The concept of ' underfunctioning ' has perhaps had its day (Vernon, P. E., 1958) in so far as it arose from a mistaken conception of what intelligence tests measure. On the other hand, it is still widely considered to have clinical value as an indication of maladjustment and it must always be important to explore the reading disability of otherwise ' intelligent ' children.

In the present sample, six children showed reading scores which differed by 2σ or more from either the vocabulary or non-verbal intelligence scores. A study of these children's performances on the main test items, along with the home-stimulation rating, were made. No clear pattern emerged and the only conclusion to be drawn is that ' underfunctioning ' has a complex causation, but the emotional attitudes of the children, even when not amounting to maladjustment, certainly seem to have considerable importance. Reading, after all, as Russell (1954) reminds us, is not a purely intellectual function, there are also ' emotional and attitudinal ' factors. Finally, the sort of association that was found between reading backwardness and poor speech does not occur in the ' underfunctioning ' cases, all of whom are competent speakers.

V.—Conclusions.

(1) A positive correlation was found between reading skill and speech development in the language aspect.

(2) The relation of reading skill to the mechanics of speech was not clear, since poor speech tended to occur in children who were also of low I.Q. and/or from defective homes.

(3) Good reading was not incompatible with maladjustment.

(4) ' Underfunctioning ' in reading had an ' emotional and attitudinal ' component, and was not associated with defective speech and language development.

Acknowledgement.—I wish to thank Miss Janet Craig (Social Worker) and the many parents, head teachers and class teachers, who so generously co-operated in providing data for this paper.

VI.—References.

Betts, E. A. (1954). Unsolved problems of reading. *Elementary English.*, Oct., 325-9.
Daniels, J. C., and Diack, H. (1958). *The Standard Reading Tests.* Chatto and Windus.
Gates, A. I. (1954). Unsolved problems of reading. *Elementary English*, Oct., 331-4.
Inglis, W. B. (1948). The early stages of reading. *Studies in Reading*, 1, London.
McLaren, V. M. (1950). Socio-economic status and reading ability. *Studies in Reading*, 2, London.
Monroe, M. (1948). *Children Who Cannot Read.* Chicago.
Morley, M. E. (1957). *Development and Disorders of Speech in Childhood.* Livingstone.
Morris, J. M. (1959). *Reading in the Primary School.* Newnes.
Neale, M. D. (1958). *Analysis of Reading Ability : Manual of Directions and Norms.* Macmillan.
Raven, J. C. (1947). *Guide to Progressive Matrices.* Lewis and Harrap.
Raven, J. C. (1958). *Guide to Using the Mill Hill Scales.* Grieve and Lewis.
Russell, D. H. (1954). Unsolved problems of reading. *Elementary English.* Oct., 335.
Sampson, O. C. (1956). A study of speech development in children of 18-30 months. *Brit. J. Educ. Psychol.*, 26 194-201.
Sampson, O. C. (1959). The speech and language development of 5-year-old children. *Brit. J. Educ. Psychol.*, 29, 217-222.
Schonell, F. J. (1942). *Backwardness in the Basic Subjects.* Oliver and Boyd.
Stott, D. H. (1958). *The Social Adjustment of Children.* London.
Vernon, M. D. (1958). *Backwardness in Reading.* Cambridge.
Vernon, P. E., (1958). A new look at intelligence testing. *Ed. Research*, 1, 1.

(*Manuscript received, May 6th,* 1916.)

FACTORS INVOLVED IN LEARNING A SIMPLE MUSICAL THEME

By G. E. R. BURROUGHS
(*Department of Education, University of Birmingham*)
AND
J. N. MORRIS
(*Willenhall Comprehensive School*)

Summary. 100 boys and girls of 13+ years of age were given repeated trials on a simple musical theme. The performance on eight consecutive trials was related to performance on a battery of musical and other tests and the factors underlying performance were elicited. It was found that the task was complex and changed in complexity during the eight trials. The learning curve for a hypothetical performer was broken down into a series of other curves corresponding to the factors revealed in the earlier part of the study. It was found that these curves did not follow the overall learning curve although they behaved in a generally regular manner.

I.—Introduction.

Although factorial studies of extended practice on printed tests by Green (1943) and Woodrow (1939) and on a single psychomotor test by Fleishman (1953) and Fleishman and Hempel (1954, 1955) have indicated variations in the factor patterns and in the weights of particular factors at various stages of practice, the application of this technique generally has been very limited; as far as can be discovered, it has never hitherto been applied to the learning of a musical item. There have been many studies of musical aptitude and musical ability and a number of tests purporting to measure fundamental musical traits and various musical skills have been developed, notably by Seashore (1919), Kwalwasser (1927), Semeonoff (1940), Lundin (1949) and others. In particular, the work of Wing (1948) and Mainwaring (1947) is well-known in this country. But the acquisition of musical skill studied dynamically has received no attention.

This paper is concerned with an experiment carried out with 100 boys and girls in their third year in the school of the second author. The experiment itself was in two parts , in the first part the children took a battery of some of the established musical tests plus an intelligence test and a musical interest test. This battery, given on a group basis, is referred to as the battery of reference tests. In the second part of the experiment, carried out on an individual basis, each child was presented with the same musical theme eight times over, he being required to sing it back after each presentation. This task is referred to as the criterion task. All these tests and tasks were scored and from the overall correlation matrix which resulted the underlying factors were calculated and the complexity of the criterion task at different stages of practice was determined.

II.—The battery of reference tests.

As the invesitigation of musical ability is still somewhat limited, and as none of the techniques of factorial studies have hitherto been applied to learning in the musical field, the choice of reference tests in this investigation was perforce both restricted and somewhat arbitrary. It is known that musical ability is a complex and not a simple unitarily functioning ability, but its components are by no means wholly determined.

Since the earliest attempt by Seashore, in 1919, to produce a standardised series of tests of musical ability (revised by the author in 1939) efforts have been made to measure this ability by testing perceptual efficiency, by using tests of preference, by using tests which endeavoured to include aesthetic or appreciative factors, by testing interpretive and creative abilities, by testing performance and by testing musical knowledge. That a number of independent factors is operating is clear, but their number and nature is not so clear. At present, therefore, a battery of reference tests is not likely to do more than sample these factors and inevitably there will be factors entering into the performance of a musical task which are elusive.

In the present experiment it was decided to use six out of seven of Wing's standardised tests of musical intelligence together with Mainwaring's test of immediate recall, a musical interest questionnaire—constructed for this experiment—and the Southend intelligence test.

A full account of Wing's tests has been given elsewhere (Wing, 1948). The result of many years of research, they are the tests most widely used and accepted in this country. The seven tests of Wing's final version fall into two sections; the first three, viz., (i) detecting the number of notes in a chord, (ii) detecting a change of a single note in a repeated chord, and (iii) detecting a change of note in a short repeated phrase, all deal with ear activity. Wing describes them as being " a kind of mental musical arithmetic." The remaining four deal with taste or preference. In these tests a good version is compared with a mutilated or weaker version and the candidate is asked to state his preference. The changes are in (iv) rhythmic accent, (v) harmony, (vi) intensity and (vii) phrasing. This final test, judging the better phrasing of two performances of a tune was omitted in the present experiment as it seemed to have little relevance to the criterion task.

Wing believed that his memory test (iii above), itself a modification of Seashore's original test, was the best group test of memory. Mainwaring's test (1931) of immediate recall of auditory experience seemed, however, to be of value in the present experiment because of its approximation to the nature of the criterion task. In both cases the child is required to listen to and to recall a tune. In the criterion task he is asked to sing it back; in Mainwaring's test he is asked to keep it in mind for a few seconds and then to answer questions about it. The answers must be educed from the recall of the tune because the questions are not known until after the tune has been played and ten seconds have elapsed.

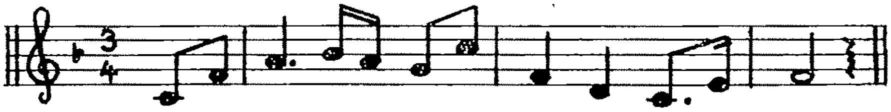
It does not follow that the ability to perceive musical material, that is to perceive rhythm or to distinguish pitch, implies any tendencies to find the material interesting. A child may score highly on Wing's first three tests yet be quite indifferent to music. Again, he may see the learning of a musical theme either as a task having musical interest or as a rote learning task, and there may be no way of distinguishing which of these two processes produces a given result. The final test of the reference battery was a musical interest questionnaire. This covered preferences, attitudes, opportunities, but not knowledge in the field. The full test can be obtained from the authors.

III.—THE CRITERION TASK.

The construction of the musical learning task was crucial to this investigation and the various considerations and stages of its development are here briefly outlined.

The first objective was to construct a tune which could be learnt by the average pupil in a reasonable number of trials. Experimental work on musical ability has shown that there is a low correlation between pitch discrimination and rhythm perception. Mainwaring (1931) obtained correlations over a series of experiments which averaged ·28, and similar results were obtained by Wing. Now if the melodic line of the proposed test was more difficult to grasp than the rhythm, the child whose pitch discrimination was superior to his rhythm perception would be in an advantageous position. The converse would result from a rhythmically difficult task. The task needed, therefore, to be of approximately equal difficulty in both its melodic and its rhythmic aspects. This could only be achieved empirically by repeated experiments with different tunes.

A second problem was to make the tune sufficiently but not too difficult. What appears at first sight to be a very simple tune, is usually found to be of considerable difficulty when successvive hearings and vocalisings are the only methods of learning it. Its reduction even to a straigthforward four-bar phrase still carried difficulties with it, for in this case it was almost always heard as two two-bar phrases. This was revealed by frequent requests for a repetition of either the first part or the second part alone. A single two-bar phrase was either too easy or imposed an artificial concentration of difficulties which made the phrase strange and even unmusical. Finally, it was found on preliminary trials that the following three-bar phrase in 3/4 time provided a sufficiently difficult yet musically satisfying theme with a reasonable balance of difficulties melodically and rhythmically.



In the main experiment the results showed that even this was a little difficult, only five of the total sample of 100 eventually giving an all-correct rendering. Nevertheless, many children succeeded in getting either rhythm or pitch quite correct and the test did give a sufficiently wide range of scores.

IV.—Experimental procedure.

As previously mentioned, the criterion task was given to the children individually. Initial experiments had shown that children tended to learn more quickly and more accurately when the tune was played on the piano than when it was sung or played on the violin. The percussive nature of the piano makes the rhythm of the music clearer than a vocal line does. Each child, then, had the nature of the task explained to him, did some preliminary lah-ing to warm up and to establish a key suitable for his voice and then had the tune played over once. His attempt to repeat it was recorded in musical shorthand and scored. The tune was then played a second time and a second attempt to sing it was made and scored, and so on for eight trials. No comment beyond that of general encouragement was made. The number of trials was limited to eight because the whole of the criterion testing programme was found to be demanding and exhausting to the pupils. Various inhibitory factors, fatigue, satiation, loss of confidence when the pupil knew he was failing but could not locate the cause of his failure, distraction and possibly others, all diminished the subjects readiness after about eight trials.

In scoring the test, one mark was given for the time value and one for the pitch of each correct note. The two semiquavers in bar one were regarded as inseparable and treated as one. This gave a total of 22 marks. Over and above this a bonus of ±20 per cent. of the marks scored up to the point where the first break-down occurred was added to meet various contingencies. The maximum additional marks available was ±4. These contingencies included singing correctly but in a key different from that in which the tune was presented, a repetition of the theme smoothly or with intermediate hesitations, and so on The range of marks was thus 0–26. The whole of this range was used.

A proportion of the responses were tape-recorded and re-scored after an interval of a week to ensure adequate reliability. A further note on this topic appears in Appendix 2.

V.—RESULTS.

(a) *The learning curve.*

That learning was taking place can be seen from Table 1, where means and standard deviations at successive trails are given.

TABLE 1

MEANS AND S.D'S OVER EIGHT TRIALS OF THE CRITERION TASK.

Trial	1	2	3	4	5	6	7	8
$\bar{x}$	6·04	8·08	9·22	9·72	10·53	11·03	11·73	11·75
σ	3·89	4·18	5·01	5·05	5·33	5·41	5·92	6·15

The means show typical learning features ; the standard deviations indicate some asymmetry of distribution and an increasing gap between the best pupil and the worst as testing proceeded. But the most striking feature is the general lowness of the scores. Even after eight trials, the 'average' individual of thirteen years managed to sing less than half of a short three-bar tune of twelve notes accurately. This agrees with Mainwaring (1933) who remarks "how limited is the ability to recall accurately any but the shortest and simplest of passages which are only presented auditorily."

A study of the individual scores as distinct from the group scores revealed two things ; first, that repetition did not produce regular and continuous improvement. 87 per cent. of the sample deteriorated at one or more of the eight stages. Learning by repetition alone, therefore, only produced spasmodic improvement. Secondly, wide individual differences at each stage of learning were seen. Not only was improvement in performance at each stage extremely irregular ; a surprisingly large number of children reached their peak at a relatively early stage. 48 per cent. of the sample scored their highest mark at stage 5 or earlier and many children deteriorated considerably after their early peaks.

Two reasons for this erratic progress were very clear. Firstly, many children trying their hardest to make some attempt to sing the whole tune through after a single hearing, sang some errors. Instead of rectifying these errors at stage 2 or subsequently, they perpetuated them. It was not what the child heard but what he thought he heard which stuck, a very old finding but still periodically

243451

forgotten. The second reason for the limited progress of some children was that while they recognised a difference between what they were singing and what they heard, they could not isolate the note(s) which was hindering progress and could not, therefore, deal with the difficulty. This emphasises the importance of directing attention to faults and weaknesses and the inefficiency of practice without guidance.

(*b*) *The factor analysis.*

The intercorrelations between the seventeen tests (eight criterion plus nine reference) were subject to a principal components analysis* with unities in the diagonals. The first eight of the obtained factors accounting for 87 per cent. of the total variance, were rotated orthogonally, the rotated loadings being given in Appendix I. The labelling of these factors is briefly indicated ·

Factor I. This is concerned with pitch, having significant loadings on those reference tests (9, 10, 11, 12) in which pitch is the major feature. As its two highest loadings are also concerned with memory and the immediate recall of auditory experience, the factor is labelled ' Memory for Melody.' It has high loadings at all stages of practice, with a tendency to decline.

Factor II. This is identified as ' Recognition of Musical Shape.' It has high loadings on intensity and harmony and is thought to be concerned with the implied harmonic structure of the theme. A simple melody has harmonic content in two ways ; first, consecutive notes have a harmonic relationship which helps in the retention of the melody (e.g., the tonic triad in inverted forms appears in the first three notes of the melody and again in the successive notes nine to twelve) ; secondly, the general run of the melody suggests or implies some form of harmonisation, often in terms of the three primary triads, which with some people assists in their learning of the theme. This factor appears throughout the criterion task, showing a steady increase in loadings.

Factor III. This factor has high loadings in verbal intelligence and interest and appears largely to be a non-musical factor which is contributing, albeit somewhat erratically, to the criterion task. The contribution tends to decrease after the early trials, as we might expect if the task is seen as one which initially is capable of a limited solution by the application of intelligence as well as musical ability.

Factor IV. This factor can be identified from its high loading in the rhythmic accent test. It is the ability to appreciate the time elements of the music (as contrasted with the melodic elements of Factor I). This factor is not sampled in the early stages of practice but builds up to influence performance later. This suggests that in learning the task the acquisition of melody preceded to some extent the acquisition of the rhythmic and time elements.

The remaining factors appear to be of very little importance in the present investigation in as much as they contributed almost nothing to the learning of the given theme. They are dealt with in detail elsewhere (Morris, 1960).

Two further points arise from the rotated factor loadings. The first is that no factor specific to the task† itself has been revealed. This contrasts with the consistently reported ' specific complex co-ordination factor ' in each of

---

* This analysis was carried out for us by Mr. W. L. B. Nixon of the University of London Computer Unit.

† A factor specific to the task is one which has significant loadings in the criterion task, usually over all the trials, but no significant loadings in any test of the reference battery. See Factor I in Fleishman and Hempel (1954) for an example.

Fleishman's studies. No variation in rotational procedures adopted in the present study succeeded in isolating such a specific musical task factor. The second point to note is that, in factorial terms, the musical task has not become less complex as practice has continued. Common sense suggests that as a task is mastered it might well become less complex and this is consonant with Fleishman's findings. In the present case it seems that the task was inadequately mastered. Although, as mentioned above, eight trials were thought to be enough, the factorial findings and the mean scores at each trial both suggest that this was wrong and that most children had by no means brought the material to be learned adequately under control. The figures, however, take no account of the observed fact that the children had had enough of it.

VI.—THE LEARNING PROCESS.

It is proposed in this section to deal more fully with the contribution made by the factor analysis to our understanding of the learning process. It is clear from the mean scores obtained from trial to trial that learning has taken place and that these figures give rise to the usual ascending but negatively accelerating learning curve. This learning curve can be thought of as the linear sum of a number of other curves, just as the curve corresponding, say, to a complex periodic function can be considered as the sum of a number of simpler periodic functions. If, in the present case, these component curves were, like the learning curve, sensibly regular, it might be postulated that these corresponded to learning processes, contributing to but more fundamental than the overall learning process. The identification of these basic processes and a study of their general form as revealed by their curves might then have important implications for learning theory. Three simple examples may help to make this clear.

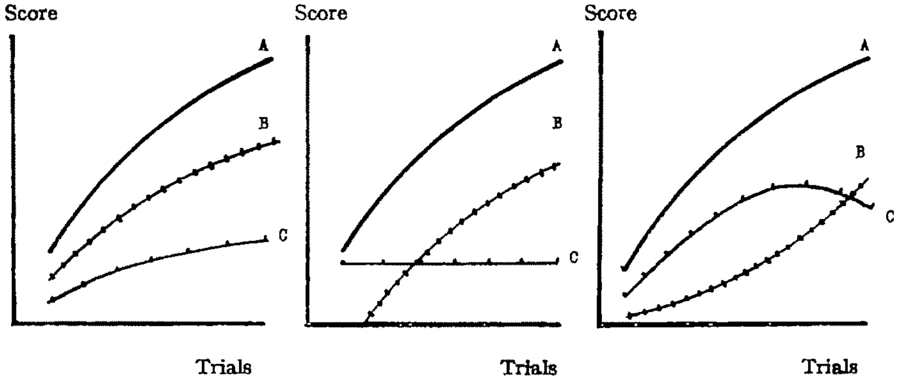


FIG. 1a FIG. 1b FIG. 1c

In each of the figures 1a, 1b and 1c, curve A is the same and is the overall learning curve, i.e., the line joining the means at consecutive trials. The figures show different ways in which curve A could be compounded linearly from two other curves, B and C, which correspond to two more fundamental abilities. In figure 1a these two curves are similar to curve A (increasing from trial to trial, but negatively accelerating). In Figure 1b curve B is similar in form to A, though starting at a later trial, but C is horizontal, i.e., its contribution to learning is constant throughout. A curve such as C in this example might indicate single trial learning. Figure 1c shows a case where one curve, B, is increasing and positively accelerating, i.e., it is contributing proportionally more and more at each trial, but curve C contributes increasingly only up to the halfway stage and thereafter falls away.

Now the basic theorem of factor analysis is that an obtained score, albeit in standardised form, is the linear sum of certain more fundamental scores; a factor analysis of a repeated task, therefore, provides the basic information required for an investigation along the lines of the previous paragraph. In particular, the principal components method of analysis allows us to calculate exactly the component scores for any real or postulated individual within the limits of the number of factors extracted. There are a number of practical difficulties to be dealt with which are considered in Appendix 2 and one obvious one has forced us at this stage to consider the performance not of the 'average' individual but of the person whose performance in the trials is consistently 1 S.D. above the mean. For such a person the smoothed curves for the first four factors are shown in Figure 2.

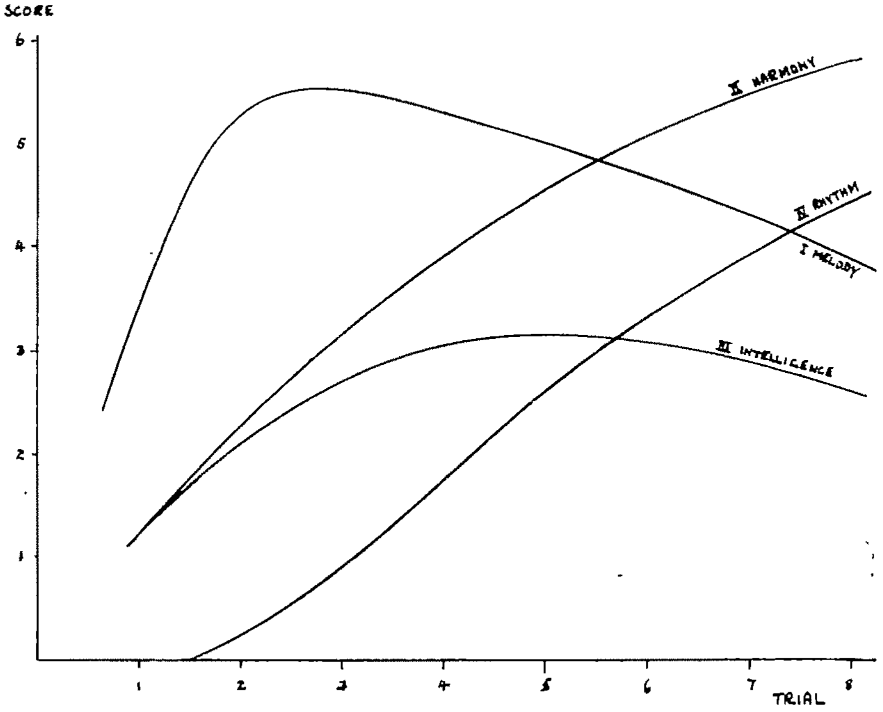


FIGURE 2

The overall learning curve, rising to a maximum of 17·90 marks at the last trial, goes too far off this figure to be shown and the remaining factors contribute less than 1 mark to the total performance at any stage.* The four curves show how many of the obtained marks at each trial, for the $+1\ \sigma$ person, can be ascribed to each of the four abilities identified in the factor analysis and their general form over the eight trials.

* The full figures, in cyclostyled form, are available from the authors.

We see :

(1) The dominance of the memory for melody factor in the early stages of learning. At its peak in the third trial it accounts for 6 of the 14 marks obtained by the $+1\ \sigma$ person.

(2) That both this factor and the factor of intelligence fall away after the half-way point.

(3) That the harmonic factor labelled " recognition of musical shape " and the rhythmic factor make increasing contribution to overall performance throughout. At the eighth trial these two factors are together responsible for approximately 11·5 of the 17·9 obtained marks.

(4) That no factor of any significance makes a uniform contribution throughout.

It is important to show the extent to which the curves in Figure 2 do sum to the original obtained curve. In Table 2 are shown the actual scores obtained by the $+1\ \sigma$ person, the scores resulting from a linear sum of the four curves and the scores found by summing all eight of the rotated factors previously mentioned. In noting the differences between the first and third columns, it will be remembered that, as required by a full principal components analysis with unities in the diagonals, there are still nine small factors which have not been considered at all.

TABLE 2

EXTENT OF AGREEMENT BETWEEN OBTAINED AND CALCULATED SCORES ON CRITERION TASK.

Trial	Mean score of $+1\ \sigma$ person	Sum of scores of four largest factors	Sum of scores of all eight factors
1	9·93	6·18	7·24
2	12·26	10·14	10·80
3	14·23	12·30	12·65
4	14·77	12·66	13·18
5	15·86	15·40	15·70
6	16·44	15·48	15·91
7	17·65	16·61	16·48
8	17·90	16·53	16·45

The figures in Table 2 strongly suggest that the four factors discussed account for most of the performance at each trial except for the earliest and especially the first trial. It seems that there are a number of factors operating as the task is attempted for the first time which the present investigation has failed to uncover. This is almost certainly due to inadequacies in the battery of reference tests. It is by no means certain that they are all of a musical nature, inadequate though the reference battery may have been in this respect. Some of the forces operating initially may be associated with set, with motivation, with personality and with other variables not easy to control. Failure to reveal them might be the cause of the apparent increase of factorial complexity with practice already commented on.

VII.—BIBLIOGRAPHY.

FLEISHMAN, E. A. (1953). Testing for psychomotor abilities by means of apparatus tests. *Psychol., Bull.*, 50, 241-262.

FLEISHMAN, E. A., and HEMPEL, W. E. (1954). Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika*, 19, 239-252.

FLEISHMAN, E. A., and HEMPEL, W. E. (1955). The relation between abilities and improvement with practice in a visual discriminating reaction task. *J. Exp. Psychol.*, 49, 5, 301-312.

GREEN, E. B. (1943). An analysis of random and systematic changes with practice. *Psychometrika* 8, 37-52.

GUTTMAN, L. (1955). A generalised simplex for factor analysis. *Psychometrika*, 20, 173-192.

HUMPHREYS, L. G. (1960). Investigations of the simplex. *Psychometrika*, 25, 313-323.

KWALWASSER, J. (1927). *Kwalwasser Test of Musical Information and Appreciation.* Iowa City Bureau of Educational Research, Univ. of Iowa.

LUNDIN, R. W. (1949). The development and validation of a set of musical ability tests. *Psychological Monographs*, No. 305.

MAINWARING, J. (1931). Experiments on the analysis of cognitive processes involved in musical ability. *Brit. J. educ. Psychol.*, 1, 182.

MAINWARING, J. (1933). Kinaesthetic factors in the recall of musical experience. *Brit. J. Psychol. (General)*, 23, 284-307.

MAINWARING, J. (1947). The assessment of musical ability. *Brit. J. educ. Psychol*, 17, 83-96.

MORRIS, J. N. (1960). *An Investigation into the Factors Involved in a Learning Task in Music.* Unpublished M.A. thesis, University of Birmingham.

SEASHORE, C. E. (1939). Psychology of music. Revision of the Seashore measures of musical talent. *Musical Educators J.*, 26, 31-33.

SEMEONOFF, B. (1940). A new approach to the testing of musical ability. *Brit. J. Psychol.*, 30, 326-340.

THOMSON, G. (1951). *The Factorial Analysis of Human Ability.* 5th Edn., Univ. of London Press, 115-117.

WING, H. (1948). Tests of musical ability and appreciation. *Brit. J. Psychol., Mono.. Supp.*, No. 27.

WOODROW, H. (1939). The application of factor-analysis to problems of practice. *J. Genet. Psychol.*, 21, 457-460.

APPENDIX I.

ROTATED* FACTOR LOADINGS.

Test	I	II	III	IV	V	VI	VII	VIII	h^2
1. Trial 1	59	29	48	−03	15	−05	−17	24	78
2. Trial 2	68	42	40	03	18	−06	−07	11	84
3. Trial 3	66	37	46	20	04	−04	−06	09	83
4. Trial 4	57	48	33	36	10	−02	−04	−15	83
5. Trial 5	64	50	46	42	05	02	−07	−04	94
6. Trial 6	55	48	40	44	12	08	−05	−02	92
7. Trial 7	46	53	36	52	12	04	08	−05	92
8. Trial 8	43	53	39	51	10	06	10	−12	91
9. Chord Analysis	30	−14	25	35	28	59	−16	−14	77
10. Pitch Change	45	−18	20	38	32	23	48	−20	84
11. Memory	52	02	24	−04	37	44	28	25	80
12. Rhythmic Accent	−06	−27	−10	51	60	−11	−17	48	98
13. Harmony	05	36	03	22	−16	76	04	36	91
14. Intensity	05	46	−18	−25	76	03	20	−01	93
15. Immediate Recall	46	16	−16	40	05	04	58	30	85
16. Intelligence	−17	10	47	19	64	24	28	00	84
17. Interest	10	19	53	13	−15	−11	55	43	87

* Since different workers rotate in different ways, we indicate here the principles which have guided us in the present study.

(*a*) Positive manifold.
(*b*) Simple structure.
(*c*) Splitting the big general factor found in the original principal components analysis among as many factors as possible consistent with psychological meaning.
(*d*) Seeking a factor specific to the task, i.e., a factor with significant loadings in the task but without significant loadings in the reference battery.

APPENDIX 2.

SOME TECHNICAL POINTS

1.—*Reliability of the criterion task.*

The obtained inter-correlations between performances in the eight trials approximates closely to what Guttman (1955) calls a simplex. The highest correlations are between adjacent trials and are, therefore, found in the diagonals of the correlation matrix, while correlations between other trails fall away as the distance between trials increases, i.e., they fall away towards the upper right hand and lower left hand corners of the matrix. This is a common feature of repeated trials intercorrelations In such a case the reliability of trial m is given by

$$r_{mm^1} = \frac{r_{lm} \cdot r_{mn}}{r_{ln}}$$

where l is any preceding and n is any subsequent trial. This allows a number of estimates to be made for all trials except the first and last. In the present case, fifty-six such estimates are available. The mean of these was ·90 and the lowest was ·73. Some of the obtained values are spuriously high due to departure of the correlations from the strict simplex pattern, but it would not be unreasonable to claim a reliability of the order of ·8 for the criterion task.

It may be mentioned in passing, that it is doubtful whether a matrix in simplex form should be factorised on a minimum rank basis. (Humphreys, 1960). While the use of unities in the diagonals, in our case, was not based on this consideration, it probably had some advantages

2.—*Calculation of a peron's factor scores from a rotated principal components matrix.*

Ordinarily a person's factor scores are obtained from his standardised test scores and the unrotated principal components matrix (Thomson, 1951). In this case—

$$z_i = \sum_k l_{ik} v_k \quad \dots \quad \dots \quad \dots \quad (1)$$

$$v_k = \frac{1}{\lambda_k} \sum_i l_{ik} z_i \quad \dots \quad \dots \quad \dots \quad (2)$$

where z_i = person's standardised score in the ith test
v_k = his factor score (also standardised) in the kth factor
l_{ik} = unrotated loadings matrix
λ_k = kth latent root = $\sum_i l^2_{ik}$

In the case of a rotated factor matrix this has to be modified. Suppose the rotation is given by the orthogonal transformation

$$u_s = \sum_k a_{sk} v_k \quad \dots \quad \dots \quad \dots \quad (3)$$

$$\sum_k a_{sk} a_{pk} = 1 \qquad p = s$$
$$= 0 \qquad p \neq s$$

then

$$v_k = \sum_s a_{sk} \, . \, u_s \quad \dots \quad \dots \quad \dots \quad (4)$$

Then from (1) and (4)

$$z_i = \sum_{k,s} l_{ik} a_{sk} u_s = \sum_s m_{is} u_s \quad \text{say} \quad \dots \quad \dots \quad (5)$$

where m_{is} = rotated loadings matrix

and from (2) and (3)

$$u_s = \sum_{k,i} \frac{1}{\lambda_k} l_{ik} a_{sk} z_i = \sum_i q_{is} z_i \quad \text{say} \quad \dots \quad \dots \quad (6)$$

These two equations correspond in the rotated case to equations (1) and (2) of the unrotated case. The values of z_i required in (6) can be written directly from (5).

The remaining term of (6) viz., $q_{is} = \sum_k \frac{l_{ik} \, a_{sk}}{\lambda_k}$ needs to be calculated separately

3.—*The use of the correlation matrix in learning experiments.*

The correlation coefficient has disadvantages in experiments such as the present in that the process of standardisation implicit in its calculation eliminates the systematic differences of variance and, more important, the systematic differences of means which are the subject of the investigation. In our case we have studied the factor learning curves of the person who was assumed to be average on the reference tests and + 1σ on the criterion task. This is not entirely satisfactory. The obvious alternative is to carry out a factor analysis of the raw-score-product matrix. For the criterion this does not give rise to any difficulty over metric; but this is not true of the reference tests. The solution might be found in re-factorising the criterion test matrix alone in this manner after the first factorisation had established the identity of the factors but the identification of first set of factors with the second is not straightforward. We are not yet certain which procedure it would have been best for us to have adopted.

(*Manuscript received, February 6th,* 1961.)

RESIDENTIAL TREATMENT OF MALADJUSTED CHILDREN: A STUDY OF SOME FACTORS RELATED TO PROGRESS IN ADJUSTMENT*

By I. R. J. PETRIE

(Manchester Education Authority)

SUMMARY. The main purpose of this enquiry was to study the progress in adjustment of a group of twenty-three maladjusted children in a special boarding school, and to relate improvement to a number of factors.

Progress was mainly assessed by comparing the initial and final test results in intelligence, reading and arithmetic, the Bristol Social Adjustment Guides, and projection techniques. In an independent assessment the Warden of the school classified the children into broad categories according to patterns of behaviour, severity of symptoms, and degree of improved adjustment, and comparisons were made between these groups.

The results show that, whilst the reliability and validity of the Bristol Guides in the assessment of maladjustment were moderately good, those of the projection tests were generally poor. All the techniques, except the W.I.S C. verbal scale, showed significant improvement but there was very little relationship between them, and only the improvement shown on the Bristol Guides was significantly related tc the Warden's own assessments of improvement The rate of progress of these children tends to be inconsistent, for those who made most progress in the first half of the experimental period did not do so in the second. Severity of symptoms and level of intelligence were slightly related to improved adjustment, but a more important factor seems to be improvement in parental attitudes.

I.—INTRODUCTION.

THERE are few reports of experimental work on the problems of maladjusted children and there is even less literature concerning their rehabilitation, especially when treatment is residential. In this country, Pringle (1957) has made a general survey of residential establishments and Reid and Hagan (1952) have done the same in America. Some reports of the work of individual establishments have also been written but these tend either to be descriptive and discursive rather than 'scientific' (Wills, 1947; Burn, 1956), or presented largely from the psychoanalytical viewpoint (Bloom, 1955; Dockar-Drysdale, 1958). This study, by describing the progress in adjustment of a group of maladjusted children in a particular school, follows upon the work of Bettelheim (1950, 1955) and Redl and Wineman (1951, 1952) but, unlike these investigations, use has been made of objective and projective techniques.

II.—THE AIM OF THE INVESTIGATION.

The purpose of this study was to examine a number of factors which might influence progress in adjustment and those considered include: the pattern of behaviour which the children showed; the severity of their symptoms; home backgrounds; intelligence; and chronological age. Special consideration was also given to those children who appeared to have made least progress in adjustment.

---

* This article is based on a thesis presented for the degree of Master in Education in the University of Manchester, 1960.

III.—THE SCHOOL AND CHILDREN.

The subjects of the investigation were pupils of an L.E.A. boarding school in the North of England. Accommodation is provided for seven girls between 7 and 11 years, and eighteen boys between 7 and 13 years. Children are admitted on the recommendation of the Child Guidance Clinic, and treatment mainly consists of environmental therapy.

All the children in school were selected for study provided that they could be observed for at least 12 months out of the 18 months experimental period. The resultant sample consisted of seven girls and sixteen boys between the ages of 7 and 12 years. Four of these had been in the school less than 6 months, six had been between 6 and 12 months, five between 12 and 18 months, and three had been in the school between 18 months and 2 years. The remaining five children joined the school during the first 6 months of the experimental period.

IV.—TESTS AND TECHNIQUES USED.

Progress was assessed by means of the Wechsler Intelligence Scale for Children, Schonell's attainment tests in reading and arithmetic, measures of overt behaviour and projection tests.

Overt behaviour was measured by the Bristol Guides (Stott, 1958). The Guides " Child in Residential Care " were completed by the Deputy Warden, and the procedure for their use in research, suggested by Stott, has been followed. He recommended that each indicator of maladjustment should be weighted two points, whilst milder symptoms of unsettledness, together with various miscellaneous symptoms should be ' scored ' as one point.

The projection techniques selected were: Test of Personality Adjustment (Rogers, 1931); Picture Frustration Study (Rosenzweig, 1948); and Family Relations Test (Bene and Anthony, 1957). On the Rogers' test the sum of four ' diagnostic scores ' provides an index of maladjustment, and the Group Conformity Rating (G.C.R.) serves the same purpose for the P.F.S.

V.—TESTING PROGRAMME.

Data from these tests and techniques were obtained initially in July, 1958; a further assessment by means of the Bristol Guides was made in March, 1959; and in the final testing of December, 1959, the initial testing programme was repeated. Progress in attainment was tested every term.

VI.—PROCEDURE.

A general survey of improvement was made and case histories for each of the twenty-three children were written. The Family Relations test proved difficult to express numerically and, therefore, was only used to a limited extent in the survey. Improvement was measured by comparing the initial with the final test results. However, the greater number of assessments made with the Bristol Guides enabled the progress made in the first 9 months of the experimental period to be compared with that made in the second. Whilst qualitative information could be used in individual cases, the assessment of improvement in the survey depended upon numerical ' scores,' and these, especially in the measurement of overt behaviour and projective techniques, were crude.

Certain comparisons were made within the sample and for this purpose the Warden classified the children into the following broad categories: (1) by behaviour; (*a*) U/W group, shy, unforthcoming children tending to be withdrawn; (*b*) HX group, children who were aggressive and showed undue anxiety to be accepted by adults; (*c*) 'Mixed' group, children showing a mixture of both the above patterns; (2) by severity of symptoms; (*a*) seriously distrubed, (*b*) moderately disturbed, (*c*) least disturbed; (3) by degree of improved adjustment; (*a*) most improvement, (*b*) moderate improvement, (*c*) least improvement.

The measures of overt behaviour and projection techniques have been presented in rank order and, therefore, non-parametric statistical analyses have been extensively used.

VII.—INITIAL TEST RESULTS.

(*a*) *Intelligence and Attainment.*

The majority of the children are of average intelligence (mean WISC 96·0) or are within one SD (14·78) of the mean. However, two children in the sample obtained I.Qs. above 125 and three fell below 85. It is perhaps interesting to note that the mean verbal I.Q. of 97·4 and mean performance I.Q. of 95·4 does not support the finding of Pringle and Bossio (1958) who gave the WISC to a group of deprived children and found that average performance quotients were significantly higher than average verbal quotients.

Retardation in reading (mean quotient 92·04) is less marked than in mechanical arithmetic (87·9) and problem arithmetic (83·1) and this confirms previous findings (Biggs, 1959).

(*b*) *Overt Behaviour.*

In a recent paper Stott (1960) has suggested that a 'score' of over 20 points on the Guides should be regarded as relatively serious maladjustment and, in our sample, seventeen children were rated above this score. The range was 14–61, the median 30 and the semi-interquartile range 8·5.

Some measure of the reliability of the Guides is afforded by the rank correlation of +·77 obtained between 'scores' on the Guides in July, 1958, and those obtained in March, 1959. The validity of these schedules was assessed by analysing the 'scores' obtained by the three groups, who had been graded by the Warden according to severity of symptoms, and the Kruskal-Wallis statistic (analysis of variance by ranks) shows that differences between these groups are significant at the ·02 level.

(*c*) *Projection Tests.*

The median 'score' obtained on the Rogers' was 48, with a range of 14–72, and the semi-interquartile range was 7. On the P.F.S. the median Group Conformity Rating was 55, range 40–83, and the semi-interquartile range was 10.

The rank correlation between the initial and final Rogers' was +·51 and between the two P.F.S. tests it was +·24. The validity of these techniques was measured by comparing their results with those of the Guides, and coefficients of +·20 and −·06 were obtained for the Rogers' and P.F.S., respectively. It could be argued, of course, that there is no *a priori* reason why techniques designed to assess underlying feelings and attitudes should be related to measures of overt behaviour.

VIII.—EFFECTS OF TREATMENT.

(*a*) *Intelligence and attainment.*

Table 1 summarizes the initial and final test results in respect of intelligence and attainment.

TABLE 1

IMPROVEMENT IN INTELLIGENCE AND ATTAINMENT.

	WISC Full	WISC 'V'	WISC 'P'	Reading	Mechanical Arith-metic	Problem Arith-metic
Initial Quot's:						
Mean	96·0	97·4	95·4	92·04	87·9	83·1
S.D.	14·78	14·25	14·36	21·61	12·27	17·58
Final Quot's :						
Mean	101·5	100·8	101·9	94·8	91·6	86·5
S.D.	17·89	14·67	15·86	21·43	19·07	19·45
Significance of Difference ..	t=3·48 ·01=2·83	1·76 ·05=2·08	2·26 ·05=2·08	2·40 ·05=2·07	3·18 ·01=2·82	3·08 ·01=2·84

No parallel form is available for the WISC and consequently there is a possibility of practice effects. This seems more likely to occur on performance items than on verbal, and therefore we should accept the non-significant difference on the verbal scale rather than the significant differences on the performance and full scales. The relatively greater progress made in mechanical and problem arithmetic, compared with that made in reading, again confirms previous findings.

(*b*) *Overt Behaviour.*

Table 2 presents the results obtained from the Bristol Guides in July, 1958, March, 1959, and December, 1959. These show that progress is consistently made throughout the experimental period and all comparisons are highly significant when one-tailed values are accepted.

TABLE 2

IMPROVEMENT IN OVERT BEHAVIOUR.

	Initial	March, 1959	Final
Median	30	21	16·5
Range	14—61	10—52	4—41
Q..................	8·5	7·5	6·25
Significance of Difference	Between Initial and March, 1959, T=22·5, ·005=49 Between March, 1959, and Final, T=34·5, ·005=43 Between Initial and Final, T= 3·5, ·005=55 (Wilcoxon Matched-Pairs Signed-Ranks Test).		

A striking result of this research was that there is no relationship between those children who made most progress between July, 1958, and March, 1959, and those who made most progress between March, 1959, and December 1959. The rank correlation obtained from the Bristol Guide assessments was —·02.

(*c*) *Projection Tests.*

It has been accepted that a lower ' score ' on the Rogers' test and a higher G.C.R. on the Picture Frustration Study are indicative of improved adjustment. The initial and final results, summarized in Table 3, show that progress is more marked in the P.F.S. than in the Rogers,' but, in this investigation, as was indicated previously, the P.F.S. is apparently less reliable than the Rogers'.

TABLE 3

IMPROVEMENT MEASURED BY THE ROGERS' AND P.F.S.

		Rogers'	P.F.S.
Initial Tests	Median	48	55
	Range	21—72	40—83
	Q	7	10
Final Tests	Median	45	65
	Range	29—67	46—92
	Q	5·25	7·5
Significance of Difference	. ...	T=53 ·025=52	T=34 ·02=38
(Wilcoxon Matched-Pairs Signed-Rank Test)			

IX.—IMPROVEMENT IN ADJUSTMENT OF PARTICULAR GROUPS.

Comparisons of improvement in adjustment among different groups within the sample were assessed by means of the Bristol Guide schedules. Furthermore, the improvement in adjustment was not calculated by differences in raw scores but by these differences expressed as percentages. This method was adopted to overcome, to some extent, the preponderance of delinquent-aggressive items in the Guides and, in particular, to make improvement shown by the U/W group more comparable to that shown by the HX group. Thus, if the initial and final Bristol Guide ' scores ' for an HX child were 48 and 36, and those for a U/W child 20 and 15, the improvement in both cases would be scored as 25 per cent., despite the differences in raw scores not being identical.

(*a*) *U/W, ' Mixed ' and HX Groups.*

Whilst the results are not definite, there is a marked tendency for the children in the U/W group to make better progress than those in the other groups. The value of H (Kruskal-Wallis statistic) for the Guides is 5·14 (·05=5·99).

(*b*) *Groups Classified According to Severity of Symptoms.*

One might expect the least seriously maladjusted group to make more progress than the moderately disturbed group and even more progress than the children in the most seriously maladjusted group. The differences between the groups, however, are not significant, although they are in the expected direction. (H=3·47 ; ·05=5·99).

(*c*) *Progress Related to Age.*

The difference in progress between the twelve older children (mean C.A., 10 years 11 months) and the 11 younger (mean C.A. 9 years 2 months) is not significant.

(*d*) *Progress of Boys and Girls.*

There was no difference between the improvement made by girls and boys.

(*e*) *Progress Related to Intelligence.*

It is widely believed that the more intelligent maladjusted children tend to make more improvement than the less intelligent, and this hypothesis was tested by comparing the progress made by the thirteen children with an I.Q. of 99 or above (mean I.Q. 107·6) with the ten children with I.Qs. below 99 (mean I.Q. 88·9). The difference between these groups, although in the expected direction, was not significant and a similar comparison between the upper and lower quartiles in the I.Q. rankings also yielded a non-significant result.

(*f*) *Progress Related to Improvement in Home Backgrounds.*

The P.S.W.'s rated the homes of nine children as having improved during the experimental period, whilst those of the remaining fourteen had either shown no improvement or had deteriorated to some extent. The difference in progress made by these groups in overt behaviour was tested but, although the value of 43·5 obtained from U (Mann Whitney statistic) is in the expected direction, it falls short of significance (·05=36, one-tailed).

The Family Relations test cannot be scored, but it was possible, on the basis of the qualitative analyses already made in the case histories, to assess the differences between the initial and final tests as being ' improved ' or ' not improved .' The six children who gained a rating of ' improved ' made significantly greater progress on the Guides than those rated as ' not improved ' (U=16·5, 0·25=21 one-tailed).

(*g*) *Groups Classified by Degree of Improved Adjustment.*

In the improvement in the test results among the three groups classified by the Warden according to degree of improved adjustment, there were three differences which were in the expected direction. For these the Kruskal-Wallis statistic gave the following values: arithmetic 1·86 (·05=5·99), Bristol-Guides 14·94 (significant at ·001 level), and P.F.S. 4·05 (·05=5·99). Improvement shown on the W.I.S.C. verbal scale, reading, and Rogers' did not differentiate the groups. The course grading adopted in the Family Relations test precluded a similar analysis from being made, but those children considered to have made most progress in adjustment are clearly identified.

The ' test ' results of the seven children classified as having made least progress, were cursorily examined. As would be expected from previous analyses the Family Relations test and ratings by the P.S.W.'s showed very little improvement in home backgrounds. Moreover, the case histories showed that marked parental instability was a major factor in three homes, and severe rejection in a further two. It might be thought that the home backgrounds of this group were such as to result in these children suffering some loss of feelings towards adults. This could be shown on the Guides by increased K (indifference to adults) and concomitant XC (anxiety to be accepted by other children). The percentage of K and XC is, in fact, higher for these children than the remainder in the sample, but the difference falls short of significance (U=35, ·05=30, one-tailed).

A qualitative study of the results of the Rogers' test also indicates this group's greater tendency towards feelings of family insecurity, as well as greater feelings of personal inferiority, and, in the P.F.S. responses, there are increased tendencies not to accept blame and to feel that every one is against them.

X.—THE RELATIONSHIPS AMONG THE ASSESSMENTS OF IMPROVEMENT.

The improvement shown between the initial and final Bristol Guides and that shown by other techniques was correlated by ranks, and the values obtained were: arithmetic +28, reading —·03, Rogers' —·18, and P.F.S. +·40.

XI.—CONCLUSIONS.

The results of this study are more suggestive than definitive; for not only was the period of observations short but the small sample studied is not necessarily representative of maladjusted children in similar schools. Other limitations include the lack of controls, the almost complete disregard of innate factors, and the coarse methods used in classifying the children and measuring their behaviour and attitudes. However, the main conclusions of the study may be summarized as follows.

(*a*) *The Value of the Techniques.*

The techniques and tests used proved of value both in the case histories and general survey, but whilst the reliability and validity of the Bristol Guides in the assessment of maladjustment were moderately good, those of the projection tests were generally poor.

Apart from the W.I.S.C. verbal scale, all the techniques showed significant improvement and we may tentatively conclude that, during the experimental period, the majority of the children in the sample modified their overt behaviour to some degree and, to a lesser extent, some changes in their attitudes and underlying feelings also took place. However, it is important to remember that lack of controls in this research has made it impossible to assess the extent to which spontaneous improvement has taken place. Lunzer (1960) in a follow-up after one year of twenty-four disturbed children attending ordinary primary schools tentatively concluded, on the basis of Bristol Guides, and a 5-point scale of improvement, that some degree of spontaneous recovery took place in rather more than half his cases.

Only the improvement shown by the Guides significantly differentiated the three groups which the Warden had graded by degree of improved adjustment. Non-significant differences in the same direction were also found in arithmetic, the P.F.S. and Family Relations test. An element of subjectivity is inherent in the Guides for they are, in effect, check lists based on the observations of an individual adult. Nevertheless they provide a reasonably good means of assessing the effects of treatment and,as their use depends more upon careful observations than specialised knowledge, they could well be employed more extensively. Whilst the projection tests are less suitable for general use, the ability of both the P.F.S. and Family Relations Test to measure improved adjustment seem to deserve further investigation. Unlike the Rogers' test they are designed to explore particular rather than general personality trends, and Vernon (1953) believes that projection tests which are more specific are likely to achieve greater validity.

(*b*) *Relationships Between the Assessments of Improvement.*

It seems that the rate of improvement made by these children tends to be inconsistent, for the Guides show that those who made most progress in overt behaviour in the first 9 months of the experimental period did not continue to do so in the second. Inconsistency in the rate of improved adjustment was also noted in a number of case histories, and this was particularly observed in the arithmetic profiles. It may be that maladjusted children lack the resources to make sustained improvement or that progress made in one direction tends to inhibit progress made in another. Certainly the lack of relationships among the various assessments of improvement seems to support the latter suggestion.

(*c*) *Comparisons of Progress within the Sample.*

The rate of progress of various groups within the sample was compared by means of the Guides, but none of these comparisons proved significant. Nevertheless, there was a marked tendency for the U/W group to make greater progress than the 'Mixed' or HX groups, and, to a lesser extent, the least disturbed children also tended to do better than the more seriously maladjusted groups. Neither chronological age nor sex was related at all to improved adjustment and the relationship with level of intelligence was only slight. It seems possible that intelligence can be over-rated as a contributory factor in the process of readjustment, especially when the treatment is mainly environmental.

However, in this study a child's rate of progress in adjustment is less related to intelligence than to improvement in his own home background. Improvement in the results of the Family Relations test is significantly related to improved overt behaviour and a similar analysis, based on the ratings made by the P.S.W.'s, approaches significance. A study of the home backgrounds of the seven children classified by the Warden as having made least progress in adjustment is revealing. Five of these homes, even in a sample of maladjusted-children, stand out as being particularly bad. It seems clear that therapy may fail, not necessarily because of limitations in the treatment rationale, but rather because of failure to modify adverse home conditions.

There is some slight evidence to support the belief that these children rated as having made least progress may have lost to some extent the capacity of making positive relationships with adults. No doubt with primary school children this attitude is largely inchoate, but it is indicated on the Guides by a higher percentage of K and XC items, and the Rogers' and P.F.S. provide further qualitative corroboration.

XII.—REFERENCES.

BENE, E., and ANTHONY, J. (1957). *Manual for the Family Relations Test.* London: N.F.E.R.

BETTELHEIM, B. (1950). *Love is Not Enough: The Treatment of Emotionally Disturbed Children.* Glencoe: Free Press.

BETTELHEIM, B. (1955). *Truants from Life: The Rehabilitation of Emotionally Disturbed Children.* Glencoe · Free Press.

BIGGS, J. B. (1959). The teaching of Mathematics—II: Attitudes to Arithmetic—number anxiety. *Educ. Res. Vol.* 1, No. 3, 6-21.

BLOOM, L. (1955). Some aspects of the residential psychotherapy of maladjusted or delinquent children. *Brit. J. Delinq.*, 6, 41-49.

BURN, M. (1956). *Mr. Lyward's Answer.* London: Hamilton.

DOCKAR-DRYSDALE, P. (1958). The residential treatment of 'frozen' children. *Brit. J. Delinq.*, 9, 106-116.

LUNZER, E. A. (1960). Aggressive and withdrawing children in the normal school. I—Patterns of behaviour. *Brit. J. Educ. Psychol.*, 30, 1-10.

PRINGLE, M. L. KELMER (1957). Differences between schools for the maladjusted and ordinary boarding schools. *Brit. J. Educ. Psychol.*, 27, 29-36.

PRINGLE, M. L. KELMER, and BOSSIO, V. (1958). A study of deprived children. *Vita Humana*, Vol. 1, Nos. 2 and 4.

REDL, F., and WINEMAN, D. (1951). *Children Who Hate.* Glencoe : Free Press.

REDL, F., and WINEMAN, D. (1952). *Controls from Within.* Glencoe : Free Press.

REID, J. H., and HAGAN, H. R. (1952). *Residential Treatment of Emotionally Disturbed Children.* New York : Child Welfare League.

ROGERS, C. R. (1931). *Manual of a Test of Personality Adjustment.* New York : Association Press.

ROSENZWEIG, S. (1948). The children's form of the Rosenzweig picture frustration study. *J. Psychol.*, 26, 141-191.

STOTT, D. H. (1958). *Manual to the Bristol Social Adjustment Guides.* London : U.L.P.

STOTT, D. H. (1960). The prediction of delinquency fron non-delinquent behaviour. *Brit. J. Delinq.*, 10, 195-210.

VERNON, P. E. (1953). *Personality Tests and Assessments.* London : Methuen.

WILLS, D. H. (1947). *The Barns Experiment.* London : Allen and Unwin.

(*Manuscript received, February 7th*, 1961.)

A FACTORIAL STUDY OF LEARNING A NEW NUMBER SYSTEM AND ITS RELATION TO ATTAINMENT, INTELLIGENCE AND TEMPERAMENT

BY M. E. HEBRON

(*Department of Psychology, University of Hull*)

SUMMARY. A factorial study of items in a new arithmetic learning situation suggests that, although knowledge of one system is the most important single factor in learning a new one, temperamental aspects of attention and set are also relevant. Subjects deemed suitable for the extraverted roles of exploration and amusements assimilated strange and unfamiliar facts with greater facility than those allocated to the more introverted roles of nursing and chores. The latter manipulated processes with greater effect once the facts were assimilated. Finally, there is the tentative suggestion that the two introverted roles tended to be associated with avoidance of error in two characteristic ways—those thought of as nurses were cautious at the first impact of test stimulus and those allocated to chores performed well during prolonged trials.

I.—THE PLAN OF THE ENQUIRY.

NINETY subjects of both sexes with a mean age of 12 years 4 months and a mean I.Q. of 98 were observed individually in a learning situation lasting from one-and-a-half to two hours. During this time they were taught a number system which utilised new digits and a notation based on eights instead of the familiar tens. The teaching method employed was closely similar to that commonly followed in schools and scores were obtained for ease in learning 125 items. Of these, twelve were selected for inclusion in a factor analysis on the grounds that there were differences at the ·10 level or better between the mean scores for the more and the less able pupils or between those for pupils of differing personality characteristics.

The learning situation and selection of variables.

Step 1. In order to ease the subject into the learning situation, he was asked (i) to sort the numbers 0–7 into serial order at speed, (ii) to read aloud three large numbers ranging in size from four to nine digits, and (iii) to write numbers of increasing length to dictation as a test of memory span. Of these, reading large numbers discriminated sufficiently for inclusion in the factor analysis. The numbers were : 1,394 ; 381,246 ; 165,274,389.

Step II. Eight experimental digits based on the shapes of cyrillic characters were learned in seven different presentations, ranging from matching them with the correct arabic numeral (first in serial and then in random order), and finally writing them from memory in two orders of presentation, random and serial, the last being dictated without the arabic numeral being shown. The last two stages in which the experimental digits were required to be written rather than visually recognised were significantly discriminating.

Step III. The subjects next learned to count using only the eight experimental digits which ranged in value from 0–7. Cards containing from 9–36 dots were used to give practice in counting and writing the number contained on each. For the first quantity 9, the experimental digit 1, was placed in the first and second positions normally occupied by the tens and units. This was read aloud as one-eight one. The largest number included in this section—36, was read and written ' four eight four ' (i.e., 4×8+4) having the experimental figure 4 in first and second places. The dots on all the cards except two were arranged in rows of four. The cards for 20 and 30, however, had the dots spaced randomly. In these the subjects had greater difficulty in avoiding omissions and double counting of dots while maintaining in mind the new sequence and method of counting. The only card which yielded a significant difference was that containing 30 dots randomly spaced. After counting dots on cards, numbers up to 63 were presented orally to be written in the experimental form. While still using cards the subjects were given a copy of the 8 times table with the suggestion that it might prove helpful. It was then noted whether the subject availed himself of this aid before the end of the dot counting or whether he continued counting each dot separately. This early use of the table in the new system proved a distinguishing characteristic of the groups.

Step IV. Fifteen pairs of experimental numbers, easy sums or fractions written on cards were presented in turn, and after reading them aloud, the subject was asked to guess which of each pair was the larger. He was then asked to check his answer by writing the equivalent value of each in the familiar tens system. Errors were corrected and explained by the experimenter. Three of the guesses distinguished between the groups. These were, in the order in which they appear in Tables 1 and 2 :

(i) A comparison of the fractions one-third and one-fifth each written with the appropriate experimental digit in the numerator and denominator

(ii) Seven divided by three and eight divided by two.

(iii) A comparison of fractions having the equivalence of a forty-seventh and a sixty-first written in the denominators as five eight seven and seven eight five, respectively.

Step V. Thirteen problems were presented on cards to be worked orally or in writing according to inclination. All numbers in the statements were in experimental form and the answers also were required to be presented in this form. Two attempts were allowed in the space of two minutes and when there was failure, the sum was explained. Three problems distinguished between the groups and, substituting the experimental digits for the verbally expressed quantities, they were as follows :

(i) What three figures added together make one eight six ? Two of the three figures must be the same.

(ii) What number multiplied by itself makes two eight nought ?

(iii) What is one (numerator) over one eight two (denominator) of one eight nought shillings and four pence ? Answer in pence.

This step completed the learning of the experimental system and there followed an attempt to ascertain how readily a second novel system would b6 learned. In this only five characters, designed by the subject himself to represent the numbers 0–4 were used. Counting, therefore, proceeded in multiples of 5, e.g., eleven would be expressed as two five one. After the subject had designed his new digits he attempted to write the numbers 0–25 using only these. Where failure occurred, the experimenter corrected it and expressed the number in the fives system. The number of demonstrations necessary failed to distinguish between the groups. After the counting, however, eight sums were presented using the subject's digits and the number of demonstrations needed in this exercise proved of significance.

Further variables for the factor analysis included the verbal and non-verbal scores of the Southend Test of Intelligence (Harrap), the Staffordshire Test of Arithmetic (Harrap), Silent Reading Test No. 1 (U.L.P.) and length of schooling.

TABLE 1

OBSERVED TETRACHORIC CORRELATIONS AND THEIR FOURTH RESIDUALS (TOP RIGHT) BETWEEN TEMPERAMENT, ATTAINMENT, INTELLIGENCE AND SCORES OBTAINED IN A NUMBER LEARNING SITUATION.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Counting in new system	X	08	09	–09	–16	–25	15	03	01	05	10	06	–09	08	15	–08	–05	10	–04	–02	05	–02	–15
2. Reading large numbers	25	X	–13	08	–10	07	–01	–20	14	–08	–08	–04	–06	–04	–04	04	03	01	00	16	10	09	–02
3. Use of tables in new system	03	25	X	–01	10	–10	–12	–11	03	–09	12	02	18	03	04	–08	–09	05	–04	16	10	–03	–13
4. Memory for new digits (i) ..	10	10	19	X	08	05	10	13	13	04	–07	–02	–23	–02	10	03	10	–16	–14	02	01	–12	–01
5. Memory for new digits (ii) ..	32	26	38	42	X	08	–13	–11	–12	22	14	04	02	–08	–04	–08	–02	–07	09	12	–10	–01	12
6. Guessing size (i)	–25	39	09	03	–11	X	05	–07	–28	–29	05	–11	08	17	–09	20	08	05	28	–01	–11	–17	31
7. Guessing size (reversed) (ii)	43	11	18	–28	16	01	X	25	21	–06	–03	–05	–07	–04	–09	–10	04	–06	–41	–01	16	06	13
8. Guessing size (iii)	–15	40	–26	11	–31	33	–44	X	00	–03	–03	–05	–15	00	20	35	–07	09	02	01	–01	–14	–14
9. Problem (new system) (i) ..	02	35	20	52	14	62	–36	22	X	–08	–22	–02	–14	–02	–06	17	06	11	21	–01	01	03	–17
10. Problem (new system) (ii) ..	02	48	03	31	43	54	–03	19	38	X	16	19	05	–03	07	–09	–07	–01	05	–02	–06	–01	08
11. Problem (new system) (iii) ..	30	40	–01	21	02	32	13	20	61	22	X	13	06	06	07	–20	–06	–02	–06	–07	–10	04	00
12. Problem (2nd new system) ..	17	37	06	19	09	15	–03	26	41	18	52	X	–12	–09	–08	–03	–06	02	08	06	06	07	–06
13. Arithmetic Test	14	48	23	–02	14	22	–04	16	31	50	34	52	X	21	00	10	04	–09	03	–13	–02	03	30
14. Non-verbal Intelligence Test	00	35	03	10	01	02	–06	31	18	39	31	45	71	X	03	–06	07	–14	–08	–08	–06	08	02
15. Verbal Intelligence Test	13	33	18	20	23	13	30	20	01	37	30	36	32	59	X	17	–08	01	–12	–20	–06	00	00
16. Reading Comprehension Test	20	13	–06	07	03	–16	11	49	17	18	46	26	42	56	74	X	–08	–02	–22	–23	–01	11	11
17. Somatotonia	14	08	06	28	22	–08	00	–18	23	04	16	14	10	07	03	–07	X	15	09	02	–05	09	–14
18. Nursing role	23	16	–20	–16	–20	16	26	05	–26	–03	16	11	13	25	13	13	–28	X	13	–32	18	–04	03
19. Chores role	–10	–17	–21	–23	–07	–32	17	22	00	07	–03	–10	13	31	03	13	–05	–12	X	34	–12	–07	–02
20. Exploration role (reversed)..	03	–12	07	–07	02	05	01	16	–09	05	15	02	00	21	00	03	–06	47	57	X	22	04	–06
21. Amusements role (reversed)	04	17	–28	–23	–17	00	22	25	–10	08	00	14	14	28	15	18	–10	50	31	–04	X	01	–19
22. Height-weight ratio	02	–07	–02	–07	04	18	–12	–10	06	07	00	–02	12	23	11	24	10	00	03	09	00	X	–02
23. Length of Schooling	03	15	20	03	05	34	30	13	20	07	22	10	–14	27	38	19	19	16	07	17	–12	07	X

TABLE 2

FACTORIAL ANALYSIS OF ITEMS IN A NUMBER LEARNING SITUATION TOGETHER WITH MEASURES OF ATTAINMENT, INTELLIGENCE AND TEMPERAMENT.

Variables	Unrotated Centroid Factors				Rotated Factors				
	I	II	III	IV	I	II	III	IV	h^2
1. Counting in new system	259	017	300	–252	173	–174	389	–097	221
2. Reading large numbers	625	263	–148	–210	592	286	290	–096	526
3. Use of tables in new system	138	296	272	021	318	–274	048	046	181
4. Memory for new digits (i)	221	465	091	315	502	–133	–184	262	373
5. Memory for new digits (ii)	254	402	439	181	473	–410	072	233	452
6. Guessing size (i)	330	143	–358	–088	316	397	023	–087	265
7. Guessing size (ii)	–125	109	–336	581	009	219	–521	398	478
8. Guessing size (iii)	276	–124	–546	091	072	620	–035	084	398
9. Problem (new system) (i)	484	612	–288	074	780	257	–150	011	697
10. Problem (new system) (ii)	582	220	–125	169	531	262	135	249	431
11. Problem (new system) (iii)	645	126	–050	–182	497	246	400	–014	468
12. Problems, second new system	556	138	–141	–023	451	290	232	083	349
13. Arithmetic Test	634	096	–245	162	466	423	180	262	497
14. Non-verbal Intelligence Test	728	–315	–067	346	200	419	440	587	753
15. Verbal Intelligence Test	678	–238	415	062	230	–047	695	393	692
16. Reading Comprehension Test	567	–393	175	272	039	169	510	539	581
17. Somatotonia	124	270	209	031	289	–216	021	044	133
18. Nursing	202	–408	–154	–520	–197	328	482	–349	500
19. Chores	077	–522	–128	456	–364	292	059	530	502
20. Exploration	–211	337	078	–072	136	–244	–236	–189	169
21. Amusements	–173	404	282	148	212	–429	–251	046	294
22. Height/weight ratio	116	–051	031	156	031	030	046	194	042
23. Length of schooling	377	–145	225	–050	118	–017	426	144	217
Variance	4·093	2·172	1·535	1·418	3·047	2·114	2·301	1·746	9·218

On the side of personality, temperament measures were afforded by the teachers' assessments for Somatotonia (Sheldon, 1942), the height-weight ratio (1940), and also by the Camp on the Island device (Highfield, 1953). A selection was made from the fifteen traits in Sheldon's list because they were of the kind easily observed in school situations. Using his numbers, those selected were: (1) assertive posture, (3) energetic, (7) bold direct manner, (13) unrestrained. Similar scores for viscerotonia and cerebrotonia failed to distinguish between the groups.

The Camp on the Island device is sociometric in that the subjects are deemed as suitable for four camp roles by the votes of their peers and companions who have known them for a year. The two roles of Exploration and Amusement tend to be allocated to quick fluent and outgoing subjects while the habitually slow, placid, dour and/or retreating are more frequently entered for the roles of nursing and chores. By far the most popular role for self-allocation is exploration. Similarly, there is a tendency to choose peers with high Exploration scores as companions in working situations. During the discussion of factors an operational definition of these four variables will be maintained, but in the discussion of a more fundamental nature which concludes the article, the assumption will be made that the two pairs of roles provide an extravert-introvert dichotomy.

II.—THE ANALYSIS.

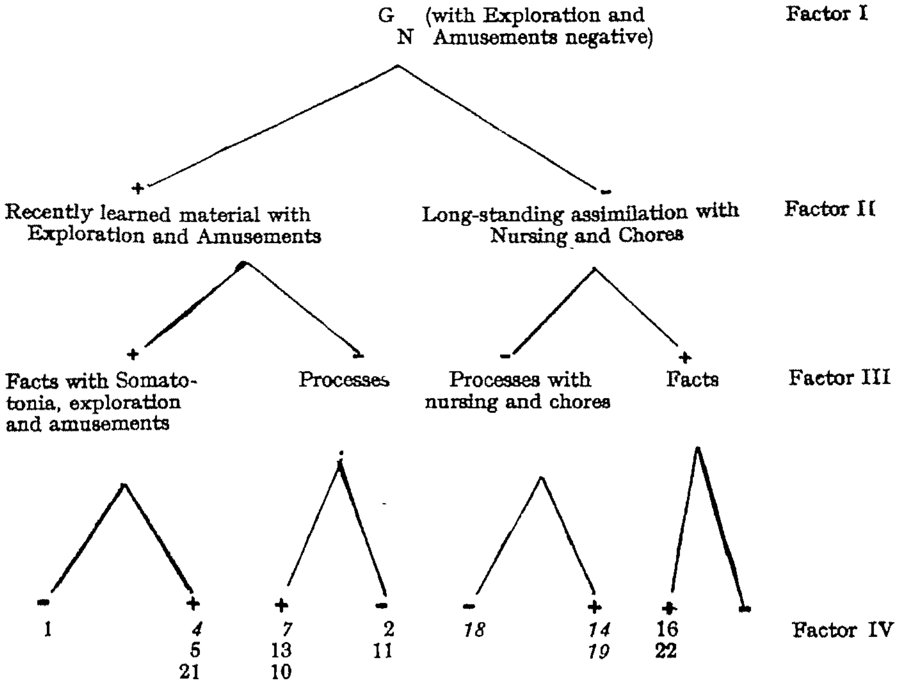
As some of the items yielded only a two-fold measure, tetrachoric correlations were calculated for all, using Thurstone's graphs. These are given in Table 1, with variables 7, 20 and 21 reversed. Coefficients of ·50 are twice their standard error. The factors were calculated to four decimal places and rounded to three in Table 2. Here the signs for the reflected variables are in their original direction

The Factors.

The first three of the unrotated and rotated factors are significant. The fourth in each case is not statistically significant but affords some clues for tentative interpretation. This is presented in Figure 1, which shows the hierarchical classification of the variables in the unrotated factors. Here, variables which have loadings of less than ·1 in the fourth factor have been omitted, while those of ·3 or more have been italicised. The remainder, which lie between these two levels, are also given.

FIGURE 1

THE UNROTATED FACTORS.


G N (with Exploration and Amusements negative)
Factor I
+
Recently learned material with Exploration and Amusements
-
Long-standing assimilation with Nursing and Chores
Factor II
+
Facts with Somatotonia, exploration and amusements
-
Processes
-
Processes with nursing and chores
+
Facts
Factor III
-
1
+
4
5
21
+
7
13
10
-
2
11
-
18
+
14
19
+
16
22
-
Factor IV

Discussion of the Factors.

The hierarchy of Figure 1 has the general factor at its apex. It can be described as indicating g activity with N-type symbols. Exploration and Amusements have negative loadings on this factor which suggests that subjects of these types tend to be at a slight disadvantage in performance of this kind.

The second factor has its higher positive loadings for reproduction in writing of the experimental digits. Correct recognition of these will also underlie success in the experimental problems and these too appear on this side of the factor. The negative loadings are for variables calling for insight into familiar symbolic forms such as the intelligence tests and the reading test. Somatotonia and the roles of exploration and amusements are associated with recall of recently learned material while nursing and chores are alligned with the more familiar forms.

The third factor presents a four-fold classification which is least ambiguous in the way it bifurcates the positive half of the second factor. This sub-divides into the elementary facts of the new number system in its positive half, with the processes arranged under the negative signs. Somatotonia, exploration and amusements appear with the facts. The bifurcation of the negative signs of the second factor results in a classification by positive sign of the two verbal tests. Since the positive sign indicates facts rather than processes in the newly learned material, it may be the element of reading vocabulary which has given rise to the common variance. Length of schooling is associated to some small extent with the facts rather than their use.

Factor IV subdivides the facts of the previous factor into the perceptual (counting dots) and the symbolic (Nos. 15 and 16); the processes of Factor III on the positive side in this factor are more complex (Nos. 7 and 14) than on the negative.

The Rotated Factors.

Rotation was effected to emphasise the characteristics indicated in Fig. I. On the cognitive side Factor 1 remains one of Ng with the newly learned material thrown into greater relief. In consequence, the negative loadings for exploration and amusements have become positive and those for nursing and chores appear as negative.

The second rotated factor contrasts processes on its positive side with the memorising of new facts on the negative. The two roles of amusements and exploration together with somatotonia appear on the factual side and nursing and chores with the processes.

The third factor is clearly of a verbal nature on its positive side and the counting of dots appears among these variables, perhaps because it calls for the same distribution of attention. Apart from the counting variables, these positive signs align activities which employ long assimilated skills. Length of schooling has most of its variance in this factor which suggests that experience has contributed to performance in the tests. These positive loadings include the roles of nursing and chores. The negative loadings are for the recently learned material together with exploration and amusements.

The fourth rotated factor has on its positive side significant or near significant loadings on the standardised tests and chores. On the negative side the variables are nursing, exploration, quantitative guessing, reading large numbers and counting. Nursing, however, has the only significant loading. This is the first factor in which the roles of chores and nursing have appeared classified by opposite signs. A tentative explanation of this is that subjects allocated to nursing are cautious in their immediate responses to stimuli and so were not misled by the size of the denominator in stating which is the larger, one third or one fifth. Those allocated more frequently to chores may have exercised obsessive care thus avoiding errors in the more prolonged activity of the tests.

III.—THE IMPLICATIONS OF THE FACTORS FOR LEARNING THEORY.

Four aspects of quantitative thinking are delineated by the factors and each of these will be considered in turn in relation to the judgments of fitness for role and the assessments for Somatotonia. This discussion will be given more point if it is accepted that the two camp roles of exploration and amusements are regarded as suited to extraverts and those of nursing and chores to introverts. Naive judges who have not observed the experimental subjects filling these roles will depend on a projection of their own feelings regarding these situations and read such adjustments of attitude into the habitual demeanours of others. There is no doubt that to fill such roles calls for temporary adjustments of attitude in the direction of extraversion or introversion and that all children experience them from the sixth year onwards in home and neighbourhood. Evidence in support of this assumption is afforded by teachers' estimates showing that outgoing expressive traits prevail among those allocated to Exploration and Amusement, the opposite picture being characteristic of those deemed by their peers to be fitted for the roles of nursing and chores (Highfield, 1953).

The theory of Eysenck (Eysenck, Franks, 1960) provides a useful interpretative tool in explaining why the camp roles should be associated with habits of attention and set in learning situations. According to this, behaviour along the extravert-introvert continuum is paralleled by a range of variation in cortical inhibition-excitation. Thus, though subdued in outward demeanour, the introvert is, nevertheless, active and sensitive when stimulated cognitively. The central inhibitory potential of the extravert, however, causes and enables him to respond to minimal clues for all practical purposes and releases affective attitudes towards them, hence the relative lack of inhibition in his demeanour. Thus, by examining the kind of mental operations which the factorial generalisations suggest, it is possible to speculate concerning the relative usefulness of introvert and extravert cognitive characteristics in performing them.

The firmest indication of association between the sociometric and ability variables is found in the second factor and suggests the hypothesis that extravert attitudes in learning favour the assimilation of the first elementary facts of the new situation while introverts are more apt when this stage is passed in applying these facts in more complex problems. It is suggested that the distinction between analytic and synthetic types is relevant here. A synthetic approach in which perceptual wholes are seen prior to their component parts would be useful in first learning to recognise the new digits and then to reproduce them in writing. The same remark applies to the stage of counting dots, where subwholes must be retained as well as the whole unstructured schema, while the separate dots are counted. An ability to afford perceptual priority to the whole would be an asset in such activities. On the other hand, in an analytic approach, in which parts are seen before the whole, there is an initial impetus towards the meaningful appraisal of a problem which will consist of the relationships between several terms. The question remains then as to whether the relatively greater cortical inhibition of the extravert predisposes him to perceive wholes while attending to a minimum of cues and similarly whether the central excitatory potential of the introvert causes him to grasp detail at the first impact.

This introvert-extravert dichotomy is also associated with recent and longstanding assimilation. Facility in the new quantitative field is most enjoyed by the extravert (Factor III) while the introvert was more at ease operating with familiar symbols of a verbal nature. This may arise from affective factors associated with this temperamental continuum.

Finally, there is a still more tentative hypothesis of two introvert approaches to error avoidance, one at the initial impact and the other in prolonged endeavour. The attitude of error avoidance is an introvert trait, but at what stage in a task it operates, and the persistence of its operation, may be a matter of habit and specific to the kind of task.

IV.—REFERENCES.

[1] EYSENCK, H. J. (Ed.) (1960). *Handbook of Abnormal Psychology*. Pitman's Medical Publishing Co., p. 461.

[2] HIGHFIELD, M. E. (1953). *Temperament in Classroom Situations*. Studies in Education, University of Hull, Vol. II, p. 72.

[3] SHELDON, W. H., and STEVENS, S. S. (1940). *The Varieties of Human Physique*. Harper Bros., p. 52.

[4] SHELDON, W. H., and STEVENS, S. S. (1942). *Varieties of Temperament*. Harper Bros., pp. 26 and 55.

(*Manuscript received, May 10th*, 1961.)

THE EFFECT OF SOCIO-ECONOMIC STATUS ON THE SCHOLASTIC ACHIEVEMENT OF SIXTH-GRADE CHILDREN

BY ROBERT L. CURRY

(*School of Education, Baylor University, Texas*)

SUMMARY. In this investigation it is shown that when a child has above-average intellectual ability he will probably overcome the effects of a deprived home environment. However, as the intellectual ability decreases, the effect of deprived social and economic conditions of the home begins to have a more serious effect on scholastic achievement.

I.—INTRODUCTION.

FOR several decades, investigations have been conducted which were designed to determine the relationship between socio-economic status and scholastic achievement. These investigations (Bryan, 1941 ; Coleman, 1940 ; Collins and Douglass, 1937 , Gough, 1946 , Shaw, 1943) have indicated the existence of a definite relationship between intellectual ability and scholastic achievement. In addition, Garrison (1932) completed a study which indicated that socio-economic status has a greater effect on scholastic achievement than intellectual ability. The common assumption has been that socio-economic status was associated with scholastic achievement regardless of the intellectual ability level.

The purpose of this study was to determine whether the differences in scholastic achievement were significant between groups of sixth-grade children when the groups were of comparable intellectual ability but differed in socio-economic status.

The subjects used in this investigation were 360 sixth-grade students who were randomly selected from 2,623 subjects from thirty-three elementary schools in a large city in the South-West United States. The subjects were Caucasians with American-born parents. In addition, they had no school record of serious emotional maladjustment and had attended the same elementary school the previous academic year.

The California Test of Mental Maturity was used to determine the intellectual ability level of the subjects, and they were assigned to one of three intellectual ability groups designated as follows : high (I.Qs. of 116 and above), medium (I.Qs. in the 94–107 range), or, low (I.Qs. of 85 and below). Each of the intellectual ability groups was further divided on the basis of socio-economic status which was determined by use of the *Questionnaire by which Socio-economic Information was Secured from Parents*, described by Eells and others (1951). Subjects were then assigned to the upper, middle, or lower socio-economic group. Each group consisted of an equal number of boys and girls to prevent any bias because of sex. Table 1 provides a distribution of subjects on the bases of socio-economic status and intellectual ability.

The California Achievement Test, Elementary Battery, was used as a measure of the scholastic achievement of the subjects. This test gives level of achievement in the areas of reading, arithmetic, language, and total achievement. Raw scores for each of the achievement areas were used for statistical analyses purposes.

TABLE 1

DISTRIBUTION OF SUBJECTS ON THE BASES OF SOCIO-ECONOMIC STATUS AND INTELLECTUAL ABILITY.

Intellectual Ability Groups	Socio-Economic Status			Totals
	Upper	Middle	Lower	
High	50	50	22	122
Medium	50	50	50	150
Low.....	14	40	34	88
Totals	114	140	106	360

Within the high, medium and low intellectual ability groups *t* tests were computed between the upper and middle, upper and lower, and the middle and lower socio-economic status groups in reading, arithmetic, language, and total achievement, A total of thirty-six *t* tests were computed with twelve being related to each of the intellectual ability groups. The ·05 level was used to determine statistical significance.

II.—RESULTS.

Table 2 provides an analysis of the data for all intellectual ability groups. In the high intellectual ability group there were no statistically significant differences in scholastic achievement between the three socio-economic groups.

TABLE 2

ANALYSIS OF DATA FOR THE THREE INTELLECTUAL ABILITY GROUPS.

Intellectual Ability Group	Achievement Area	Socio-economic Status Groups			*t*-ratios		
		Upper	Middle Means	Lower	Upper-Middle	Upper-Lower	Middle-Lower
High	Reading ..	118·1	116·2	114·9	1·04	1·34	0·58
	Arithmetic	82·6	77·7	79·8	0·92	0·95	0·70
	Language	61·3	59·3	59·7	1·45	1·02	0·21
	Total ..	262·0	253·1	254·4	1·98	1·44	0·22
Medium . .	Reading ..	100·1	98·9	96·8	0·38	1·17	0·63
	Arithmetic	66·4	66·3	63·1	0·02	1·72	0·57
	Language .	51·9	51·8	47·1	0·06	2·84*	2·49*
	Total ...	218·3	217·2	206·9	0·19	2·17*	1·66
Low	Reading ..	82·1	70·2	58·1	2·50*	4·11*	2·99*
	Arithmetic	48·1	46·6	41·7	0·37	2·06	1·66
	Language	42·4	36·2	32·9	2·20*	3·03*	1·52
	Total . .	172·7	153·1	132·8	1·99	3·55*	2·62*

*—Denotes significance at the ·05 level.

In the medium-intellectual ability group, differences in language achievement were found to be statistically significant between the upper and lower and the middle and lower socio-economic status groups. Also, the difference in the area of total achievement between the upper and lower socio-economic status groups was found to be statistically significant. Other differences in achievement were not found to be statistically significant. When differences were statistically significant the results always indicated greater achievement for the higher of the two socio-economic status groups being compared.

In the low intellectual ability group, differences in reading achievement were statistically significant between the upper and middle, upper and lower, and middle and lower socio-economic status groups. In language achievement the differences were statistically significant between the upper and middle and upper and lower socio-economic status groups. Also, in total achievement the differences were statistically significant between the upper and lower and middle and lower socio-economic groups. Other differences were not statistically significant. Significant differences also indicated greater achievement for the higher of the two socio-economic groups being compared.

III.—CONCULSIONS.

From the results of this investigation, the following conclusions were made :

(1) Socio-economic status seems to have no effect upon the scholastic achievement of sixth-grade students when the students have high intellectual ability. High-intellectual ability offsets any deficiency which may be created by lower social and economic conditions.

(2) Social and economic factors have an effect upon language achievement in the medium intellectual ability group. The upper and middle socio-economic status groups both achieve a greater amount than the lower socio-economic status group. Likewise, in total achievement, the upper socio-economic status group achieves a greater amount than the lower socio-economic status group.

(3) In the low intellectual ability group, social and economic factors have an effect on achievement in reading, language, and total achievement. In reading the upper socio-economic status group shows a greater amount of achievement than the middle and lower socio-economic status groups, and the middle socio-economic status group shows greater achievement than the lower socio-economic group. In language the upper socio-economic status group achieves more than the middle and lower socio-economic groups. In the area of total achievement, the upper socio-economic group achieves a greater amount than the lower and the middle socio-economic group achieves greater than the lower socio-economic group.

(4) As the intellectual ability decreases from high to low, the effect of social and economic conditions on scholastic achievement increases greatly.

(5) Achievement in arithmetic seems to be relatively free from the influence of social and economic conditions since no significant differences were found within any of the intellectual ability groups.

IV.—REFERENCES.

[1] ALLEN, MILDRED M. (1944). Relationship between Kuhlmann-Anderson intelligence test and academic achievement in Grade IV. *J. Educ. Psychol.*, 35, 229-39.

[2] BRYAN, RUTH (1941). *A Study of the Relationship Between Socio-Economic Status and Scholastic Achievement.* Unpublished Master's Thesis, University of Iowa.

[3] COLEMAN, HUBERT, A. (1940). The relationship of socio-economic status to the performance of junior high school students. *J. Exp. Educ.*, 9, 61-63.

[4] COLLINS, JOSEPH H., and DOUGLASS, HARL, R. (1937). The socio-economic status of home as a factor in success in the junior high schools. *Elementary School J.*, 38, 107-13.

[5] EELLS, KENNETH, *et al.* (1951). *Intelligence and Cultural Differences.* Chicago : Univ. of Chicago Press

[6] GARRISON, K. C. (1932). The relative influence of intelligence and socio-cultural status upon the information possessed by first-grade children. *J. Soc. Psychol.*, 3, 362-67.

[7] GOUGH, HARRISON, G. (1946). The relationship of socio-economic status to personality inventory and achievement test scores. *J. Educ. Psychol.*, 37, 527-40.

[8] LINE, W., and GLEN, J. S. (1935). Some relationships between intelligence and achievement in the public schools. *J. Educ. Res.*, 33, 582-87.

[9] SHAW, DUANE C. (1943). The relation of socio-economic status to educational achievement in grades four to eight. *J. Educ. Res.*, 37, 197-201.

(*Manuscript received, February 13th,* 1961.)

ABSTRACT AND CONCRETE MODES OF CLASSIFICATION IN A PRIMITIVE SOCIETY

By D. R. PRICE-WILLIAMS
(*London School of Economics*)

SUMMARY. 1.—Past investigations of the cognitive processes of primitive peoples are reviewed, with special reference to the continuum of abstract and concrete. It is pointed out that these studies have used Western type tests in reaching their conclusions. The present study differs in using indigenous material; the author, having lived among the subjects in the manner of an anthropologist, and speaking (inexpertly but sufficiently) the native language.

2.—Bush and primary school children, living in the same area, were compared in their ability to classify and to sort (*a*) models of animals known in the area, (*b*) plants actually picked in the neighbourhood. The ability to shift from one classification to another and the basis of classification were investigated. The linguistic background relevant to the classifications is described.

3.—Using familiar material there was found no difference between the two sets of children, of an age range from approximately 6½ to 11 years. The results are discussed in the light of the general ability to abstract, in the sense of shifting from one class to another, and in the framework of the transition of the process of abstraction during this age period which comes under Piaget's period of "Concrete Operations."

I.—INTRODUCTION.

THERE have been several kinds of evaluations of the cognitive abilities of underdeveloped peoples. The first kind is best exemplified by Carothers[1] who, on the basis of little relevant evidence, claimed that the African does not have sound abstraction. Also Haward and Roland,[2, 3, 4] in a series of papers in an anthropological journal, which sparked off highly critical responses by anthropologists, drew attention to what they termed the concreteness of the Nigerian's mental approach. This conclusion was inferred from the application of a somewhat inadequate (for this purpose) instrument, namely the Goodenough Draw-a-Man Test. A second kind of assessment, one which comes to the same conclusions although somewhat modified, has been that expressed by Wintringer.[5] The difference is that the latter has based his conclusions on a considerable number of relevant tests by other investigatiors. Wintringer writes:

> " L'infériorité intellectuelle du noir s'explique donc par un comportement mental profondément conditionné par une attitude concrète, intuitive, et rivée à la perception syncrétique de la réalité sensible . Par conséquent la presque totalité des noirs est incapable d'effectuer une opération mentale selon le procédé classique de l'abstraction et de manier simultanément des relations conceptuelles opposées."

On the other hand, McConnell,[6] in his study among the Tepehuan Indians, using an adaptation of the Wechsler-Bellevue version of the Kohs Block Test, found marked evidence of ability to abstract. However, both these studies are similar in that actual experiments were performed, and the conclusions drawn are not sweeping generalisations. Comparing them one might come simply to the conclusion that while Africans do not have the ability to abstract, Tepehuan Indians have.

A third kind of assessment starts off from the premise that *degree* of abstraction is the relevant factor, and has gone on to make the highly important inquiry as to whether there is a difference between acculturated and literate *versus* unacculturated and illiterate natives. Jahoda,[7] for instance, using the

Goldstein-Scheerer Cube Test on adolescent boys in or near Accra, found that "boys from literate homes performed significantly better than those from illiterate ones." Maistriaux,[8] in a whole series of investigations in the Belgian Congo, found large differences between bush and school Africans. His general conclusions are that, while qualitatively the thought processes of the African Negro are similar to those of Europeans, their rate of development in this direction is lower.

All of these studies, performed within various age ranges, but generally on adolescents and young adults, have used tests of abstract ability which are similar to those used in American or European cultures. Jahoda, in the same article, maintains that these are no more 'culture-free' than tests of intelligence, which are clearly recognised as not being so. This statement provides a starting-off point for the present study which is different from those listed above inasmuch as it employs material which is indigenous to those to whom it is applied. The question can then be directly tested, whether the difference between bush subjects and subjects with school education, still persists.

The subjects of the research were children of the Tiv tribe, of Benue Province Nigeria. The investigation was carried out while the writer was living, in the manner of an anthropologist, in a native hut of a particular compound. Experiments were carried out either in this hut, or anywhere else in the compound, or in other compounds or in the nearby Bush Junior Primary School. The study might be described as ecological, a type of research which Leeper[9] has advocated in his chapter on Cognitive Processes. Though it is recognised that properly conducted laboratory studies are methodologically more rigorous, there is much to be said for this kind of research where the instructions are carried out in the native language and where experimental results are assessed in the context of everyday living conditions. Tiv are a tribe of subsistent farmers; their interests lie in animals, plants and matters of agriculture generally. Therefore, the material used for the research were exemplars of animals and plants with which they were familiar. The purpose of the research was to see to what extent and in what kind of way Tiv children could categorise this material. The results are assessed in the framework of abstraction *versus* concreteness of approach.

As with so many psychological terms, 'abstract' and 'concrete' suffer from a lack of precision. The various definitions given to them in the literature designate two widely contrasted attitudes. To Lévy-Bruhl,[10] in his writings on the so-called pre-logical mentality, concreteness had meant dependency on particular objects, with the consequent inability to select and classify independently, attributes from the objects. To Goldstein and Scheerer[11] in their classic monograph, concreteness of thinking indicated 'immediacy of claim' of one particular aspect of an object, which might be experienced as sensory impressiveness, or as sensory cohesion, or as situational belongingness. Whatever the nature of the claim, the individual is unable to detach himself from the uniqueness of the object and fails to see it as a representative of a class. This dependency results in rigidity and 'lack of shifting.' Conversely, the abstract attitude forms the basis for being able to shift reflectively from one aspect of a situation to another or to hold in mind simultaneously various objects. Hanfmann and Kasanin[12] entertain similar definitions, while Humphrey[13] treating abstraction more as a process than an attitude, nevertheless defines it as the selection of some particular feature of a situation to the exclusion of others.

Generalising from these descriptions, abstraction will be regarded as being steps away from a 'zero-level' of complete sensory immediateness and object dependency. A simple measurement device for scoring responses on this basis will be described later.

II.—METHODOLOGY.

(1) *Assumptions.* On the assumption that (*a*) the continuum abstract-concrete has some value, particularly when considered in the setting of non-Western children undergoing Western education, and (*b*) that tests, standardised in the West, are inadequate for testing abstract ability, experiments were devised which involved problems of abstraction on 'natural' phenomena. These experiments were done with children who attended a Bush Junior Primary School and children who did not attend any school. The homes of both groups of children were identical. Both lived in compounds around the school and, with the exception of those children whose fathers were teachers (two children only), their parents were illiterate. There was one factor which could not be controlled, and that is bilingualism. Of the four classes in a Bush Primary School, the first three are taught as much in English as in their own language. Although all these experiments were carried out in the vernacular, at first with the help of translators (teachers), and afterwards by the investigator himself, 75 per cent. of the literates could speak and understand varying degrees of English. Although some of the illiterates in their turn could speak English, this was of the pidgin variety and did not go very far.

(2) *Design.* The material consisted of animals and plants. The difficulty of language (for the investigator) and the desire not to complicate the task inordinately, led to very simple methods of enquiry, which were requests for classification and requests for sorting. On the one hand, a selection of material was put before the subject with the request to select those that belonged together. This is the simplest variant of what Vinacke[14] has labelled the problem-solving method involved in concept formation experiments. On the other hand, three or four objects were put before a subject, one of which was different from the others on some basis: the subject was merely asked to say which one. In both methods reasons were sought as to the basis of classification and differentiation.

(3) *Age sampling.* This was a considerable problem as ***absolute*** age was very difficult to ascertain, but it was easier if ***relative*** age was taken as a basis. A convenient starting point for age sampling was the Bush Junior Primary School, which was divided into four classes. With the aid of its teachers, the classes were ascribed a somewhat rough estimate of age as follows:

Class I	6½ years. approx.
Class II	8 years approx.
Class III	9½ years approx.
Class IV	11 years approx.

This estimation was based on the average of each class, but was given somewhat more precision by taking out (for reasons of sampling in testing) any particular boy or girl whom the teachers knew to be clearly older or younger than the estimated average age. For example, an 11-year-old (estimated) boy who was in Class III would be relegated, for testing only, to Class IV. With this as a base-line illiterates could be classified by the simple device of asking each child whether he was older or younger than a selected boy whom he knew in school, and again whether he was older or younger than another illiterate. Generally, there could be found illiterates living side by side with literates, and although their absolute ages were always in doubt, the importance of age grading in the Tiv, and the importance of seniority in the family, determined fairly well the required comparisons.

The number of schoolchildren in the classes varied, but for convenience of tabulating the number of subjects was taken as 20 for each. This was done by leaving out the surplus around the centre of each class, again with the advice of the teachers. Collecting the same number of illiterates was a formidable task, partly on account of the continual movement of Tiv children, so the total number of illiterates equivalent to each literate class was less, namely 15. This gave a total of 80 literate subjects and 60 illiterates.

III.—LINGUISTIC BACKGROUND.

It is necessary to include a relevant digression on the classifications of animals and plants in the Tiv language, although none of the children tested overtly used them. Abraham[15] noted that there was a mixture of abstraction and concreteness in the Tiv language. The verb to break, for example, has no generic term but differs according to whether one is breaking something into pieces, or breaking it across and so forth. On the other hand, the linguistic classification of animals separates clawed animals from the cloven-hoofed and ruminants, and domestic animals from wild. Nevertheless, concreteness is also in evidence by the dependency on colour or situation: crabs are divided into red or black, fish into those that swim on the surface or those which frequent the river bed. Plants tend to be categorised functionally in the language. Bohannan[16] reports that staple crops are termed faeces-forming food while the remainder are termed urine-forming food. Abraham again comments that there is, for example, a separate classification for plants from which soup can be made.

According to my informants in the area, the linguistic classification of plants is twofold. This is not mentioned in the published literature nor did I hear of it from missionaries in the area who were familiar with the language and culture of the Tiv.

I. *Sham i yer*. This is roughly translatable as 'plants as seeds.' The term ***sham*** refers to the seed actually in the plant, while *yer* means 'to conceal.' The classification includes such plants as guinea-corn, millet, bean, ground nut, rice and benniseed.

II. *Atam a yer*. Roughly translatable as plants with roots. *Atam* is the plural of *ityamegh*, meaning fruit. This includes cassava, potatoes, and two plants, untranslatable, called *alum* and *mongolo*.

The first group has a sub-classification:

(*a*) *Ishange mom*. '*Ishange*' is the term given to a grain of seed. '*Mom*' merely means 'one.' This refers to the actual seed leaf, which is visible. It could be translatable into our own botanical nomenclature as monocotyledon. It includes millet and guinea-corn.

(*b*) *Asange ahar*. This is the plural form, and can be called 'two-leaf' or dicotyledon. It includes ground-nut, bean.

It is interesting to note that whereas this sub-classification is phenotypical, its superordinate is genotypical.

Although these classifications may be a modernised Tiv version (as I never came across other than educated Tiv who used them) the *idea*, of some plants having one seed-leaf and others two, was certainly entertained. Besides this some plants, distinguishable overtly from the language, were classified according as to whether they lay beside a stream or on a hill, and again some plants were classified in sex terms (*ihura nomso, ihura kwase*: *ihura* being the name of the plant and the other two terms being male and female). This did not imply any biological connection between the plants; it was only that each was used in a male or a female ceremony.

IV.—CLASSIFICATION EXPERIMENTS.

Procedure and Scoring. A collection of either animal examples or plant examples was arranged before each subject with instructions to put into rows those which belonged together. He was asked to do this in as many different ways as he could. A simple scoring device was to list the number of alternative ways which each subject employed, together with the reason for them. This was done on the commonly held assumption that the ability to 'shift,' that is of alternative ways of classification, is a measure of ability to abstract. The reasons were given a qualitative grading on a scale ranging from concreteness to abstraction, a zero score being a classification in terms of sensory immediateness or object-dependency, while scores above zero signified varying steps away from the particular and immediate towards generality and abstraction. As there is always an arbitrariness about such qualitative scores these results are considered as secondary in significance to the number of shifts.

(1) *Animal Material.* Toy animals were used, all familiar to Tiv children, as follows :

2 Black Cocks.
2 Brown Cocks.
2 White Hens.
2 Black Hens.
2 Brown Horses.
2 Black-and-white Cows.
2 Brown Goats.
2 Grey Elephants.*
2 Green Snakes.
2 White Sheep.
2 Brown Birds (unspecified variety).
2 Brown Rats.

In addition to this, I collected two dead black beetles and a fish known to the Tiv. The whole collection was displayed to each child at one time.

The following reasons of classification were given :

A. Individual creatures according to name.
B. Colour.
C. Size.
D. Number of Legs.
E. Situations. Example, fish in water, snake in bush, cocks and hens in compound, bird in air.
F. Domesticity *versus* Wildness.
G. Edibility.
H—Odd choices. Example, Some have marks, some are bright, etc.

Quantitative responses are given in Tables 1 and 2. As more than one classification was selected, Table 2, which shows the percentages of reasons for classifications, takes only the initial choice.

* There were no elephants in the area, but the animal was known.

TABLE 1

AVERAGE NUMBER OF SHIFTS PER AGE GROUP.

	Literates	Illiterates
Class I	3·0	3·2
Class II	3·5	3·6
Class III	4·8	5·0
Class IV	6·1	5·8

TABLE 2

PERCENTAGE OF QUALITATIVE RESPONSES PER AGE GROUP, ACCORDING TO PREVIOUS CATEGORIES, FIRST CHOICE.

	Literates								Illiterates							
	A	B	C	D	E	F	G	H	A	B	C	D	E	F	G	H
Class I	80	15	5	—	—	—	—	—	87	7	7	—	—	—	—	—
Class II	80	10	5	5	—	—	—	—	80	7	7	—	—	—	—	7
Class III	65	15	5	5	—	—	5	5	73	7	—	—	—	13	—	7
Class IV	60	5	5	5	5	—	15	5	67	7	—	—	7	13	7	—

(2) *Plant Material*. The following plants and leaves were collected from the area and displayed for classification:

guinea-corn,
millet,
benniseed,
ground-nut,
cassava,
yams,
leaves of *mongolo*,
three other types of plants, non-edible and untranslatable.

This whole collection was also displayed to each child at one time.

Qualitative responses were based on the following:

A. Edibility.

B. Size.

C. Coming from the earth or from trees.

D. Roots *versus* non Roots.

E. Situation, e.g., near river or on top of hill.

F. Seed-leaves.

Results with the plant experiments are seen in Tables 3 and 4. Table 4 again takes the initial choice of classification for the basis of scoring the reasons :

TABLE 3

AVERAGE NUMBER OF SHIFTS PER AGE GROUP.

	Literates	Illiterates
Class I	2·2	2·4
Class II	2·8	3·0
Class III	4·5	4·5
Class IV	5·4	5·6

TABLE 4

PERCENTAGE OF QUALITATIVE RESPONSES PER AGE GROUP, ACCORDING TO PREVIOUS CATEGORIES, FIRST CHOICE.

	Literates						Illiterates					
	A	B	C	D	E	F	A	B	C	D	E	F
Class I	85	5	5	—	5	—	87	7	7	—	—	—
Class II	80	5	5	—	10	—	80	13	7	—	—	—
Class III . .	70	—	10	5	10	5	73	7	7	7	7	—
Class IV .. .	65	—	20	5	5	5	67	—	13	7	7	7

V.—SORTING EXPERIMENTS.

Procedure and Scoring. The qualitative responses given by subjects on the preceding classification experiments were made the basis for sorting experiments. The procedure involved was very simple : a variety of items was put before each subject with the statement that one was different. Each subject was asked which one was different and to give the reason. A preliminary test was carried out with sizes of leaves. Four leaves were set on the ground, one of which was very big, while the other were quite small. Scoring was done in percentages of correctness of responses. A response was judged correct if both the response and the reason for it accorded with the experimenter's choice. The following list shows the material used :

(1) Animal Material.
  (*a*) Three animals with four legs, one of two legs.
  (*b*) Three domestic animals, one wild animal, as recognised by Tiv.
  (*c*) Three edible (by Tiv) animals, one non-edible.
(2) Plant and leaves Material.
  (*a*) Three plants with roots, one without.
  (*b*) Three plants with two seed leaves, one with one seed-leaf.
  (*c*) Three plant-leaves from the earth, one from a tree.

All material was selected to be as individually different from one another as possible except for the chosen criterion. For example, where model animals were used, all four objects were different colours, all were different sizes, and so forth.

Results.

Table 5 shows the percentage responses of both groups according to the material employed.

TABLE 5

PERCENTAGE OF CORRECT RESPONSES IN SORTING EXPERIMENTS

Basis	Literates				Illiterates			
	I	II	III	IV	I	II	III	IV
(1) Animals								
(*a*) Legs	75	80	90	100	60	73	93	93
(*b*) Domesticity	45	60	75	80	53	73	73	73
(*c*) Edibility	55	55	65	90	60	67	80	93
(2) Plants :								
(*a*) Roots	45	50	60	70	53	60	67	80
(*b*) Two-leaved	55	55	60	70	60	67	73	80
(*c*) From earth ...	65	70	85	95	67	73	93	100

VI.—ANALYSIS OF REASONS FOR CLASSIFICATION.

In the Introduction it was stated that sensory immediateness and object dependency formed the ' zero-level ' of the idea of concreteness. In relation to the actual responses noted in Tables 2 and 4, this means that any answer given in terms of an immediate sensory impression, such as colour, shape or size will be given a zero mark. Responses by name qualify also for a zero mark as this is evidence of object-dependency, that is the inability to conceive of the object as a member of a class. Where there is clear evidence of a movement away from the immediate situation, to make some classification on a basis of a dichotomy like edibility *versus* non-edibility, or domesticity *versus* wildness, a mark of 2 will be given. This still leaves responses concerning classifications on the basis of number of legs in animals, and roots and seed-leaves of plants to be considered. These qualify for sensory immediateness inasmuch as the exemplars are there in front of the classifier, yet there is some attempt to select from each example a feature which is common to the set, and on this basis to form a class. These responses will be given an intermediate mark of 1. The scoring system is summarised as follows :

Score O. *Animal Classifications* : Name, Colour, Size, Situation.
Plant and Leaf Classifications : Size, Coming from earth or trees, Situation.

Score 1. *Animal Classifications* : Number of legs.
Plant and Leaf Classifications : Roots *versus* non-roots, number of seed-leaves.

Score 2. *Animal Classifications* : Domesticity *versus* wildness. Edibility.
Plant and Leaf Classifications : Edibility.

Table 6 gives the comparison of the four groups of literates and illiterates in terms of the scoring system, expressed as percentages. Odd choices are not scored.

TABLE 6

SCORING OF QUALITATIVE RESPONSES.

				Animals				
Score	Lit. I	Illit. I	Lit II	Illit. II	Lit. III	Illit. III	Lit. IV	Illit. IV
0	100	100	95	94	85	80	75	80
1	—	—	5	—	5	—	5	—
2	—	—	—	—	5	13	15	20
				Plants				
0	15	13	20	20	20	20	25	20
1	—	—	—	—	10	7	10	13
2	85	87	80	80	70	73	65	67

VII.—DISCUSSION.

A highly relevant feature involved in analysing the process of classification in non-Western people is that of interest or motivation. Where there is little incentive in forming different kinds of categories, it is not surprising to find apparent dependency upon the concrete. Yet only a casual acquaintance with the monographs of anthropologists should convince one that the simplest of peoples often have social categories of kinship, marriage, property and law which call for abstraction of a fairly complex order. It is in relation to *objects* that the concrete is dominant, or appears dominant and it may be, as Rollings[17] has recently pointed out, that there is a radical difference between Europeans and Africans in relation to the perception of objects and to people. However, with reference to the present theme, it need only be said that there is little interest in forming classification of objects. Where there *is* interest, as is seen in the figures for plants in Table 6, then an abstract classification is appreciably formed. The lack of a classification on the grounds of edibility and domesticity *versus* wildness in animals needs to be explained. In the particular area in which this research was conducted, there were, in fact, very few wild animals. Most of the land had been cultivated and the game killed. Also the majority of food-stuffs came from plants. Chickens, goats and occasionally sheep were the only animal foods eaten. This factor of interest is relevant again when studying those Tables relating to shift of classifications and those in which the qualitative responses of the first choice are noted. For, in those cases where the first choice is clearly of a concrete nature, it is only by probing the individual that the interrogator receives an abstract classification. For example, with the animal material, it is only by the last or one but last shift of classification that a category of wildness *versus* domesticity comes. It does not come readily. Nevertheless, the fact that it comes at all indicates that the Tiv child can abstract. Even if the qualitative scoring system presented here is arbitrary, the number of shifts shown in Tables 1 and 3 shows that the Tiv child is quite capable of moving from one category to another. Granted that this is so, a comparison of literates and illiterates was arranged in order to see whether the small amount of formal abstraction which literate Tiv children receive in school would influence their classification of natural phenomena. A comparison of the number of shifts (Tables 1 and 3) shows hardly any difference between the groups in any of the classes. Neither does a comparison of the qualitative responses (Tables 2 and 4). The sorting experiments (Table 5) show similar results. It is difficult to comment upon this finding. The available evidence does not indicate whether the lack of difference shows a similarity of potential for abstraction, or whether the type of training in schools, up to the age tested, does not encourage an abstract attitude, or whether the type of material used for experiments fails to discriminate between the groups.

With both literates and illiterates there is a tendency for less dependence upon the concrete as the children grow older. Experiments with animals show both a decrease of the purely concrete reasons for classifications and an increase in the more abstract reasons (Table 6). An exception to this trend appears in the figures for edibility in the plants experiments. Here Class I (literates and illiterates) returned a higher percentage than the three other classes, and there is a gradual decline from Class I to Class IV, although in all groups the abstract reasons (Score 2) were the most frequent. Again, in the sorting experiments, the figures suggest (Table 5) that there is a consistent increase over the years in the ability to determine what exemplars make a group, although these figures do not show any significant difference between the kinds of exemplars chosen by the experimenter.

One point needs to be made regarding the position of language. It has been said that there was little reference to the formal system of classification reported by ethnographers. For example, the term for a domestic animal is '*ilev*.' Of the responses noted as domesticity *versus* wildness of animals, very few children actually used the term '*ilev*.' The distinction was eked out by circuitous descriptions of animals that one found in the compound and which could be left to roam about on their own, which did not attack one and the like. It was clear here that the category of ' domestic animal ' was uppermost although the actual term was not used. The findings reported, therefore, are not fully determined by the existing linguistic classification. It was also noted that linguistic terms could be constructed if, in fact, there was no corresponding term in the language. For example, there is no Tiv term for ' triangle.' When (in other perceptual experiments which the writer performed) this figure was drawn, it was labelled a ' three-cornered square.' Actually, there is really no term for square either, but a drawing of a square proper or a rectangle was called after the native term for that shaped hut—'*gondo*.' A three-cornered '*gondo*' was merely an extension, constructed to fill a lacuna in their own language That the experimenter has now introduced the *concept* of triangularity into the Tiv tribe is a philosophical point which he feels unqualified to amplify.

The discussion has so far been focussed on the rather more clinical connotation of the verb ' to abstract.' In the Introduction it was noted that the word, as far as psychologists employ it, has two aspects : that of attitude and that of process. As far as the former is concerned, we hope to have shown that there is little difference between these African children tested and European children. When we come to consider the results viewing abstraction as a process, it is necessary to use a framework other than the clinical one. It is doubtless a confusion of the two aspects that has given rise in the past to such statements as that Africans show a schizophrenic rigidity in their thinking. A convenient framework comes from the general works of Piaget. Using this framework, Maistriaux found that his adolescent and young adult African subjects, both bush and literates, only reached the most primitive of Piaget's rational stages, i.e., the period of concrete operations (see Duijker and Frijda summary[18]). Now the age range of the subjects used in the present study comes under the period of the third stage—the period of concrete operations. It is within this period, according to Piaget, that the idea of classes properly arises. As Berlyne[19] puts it: "The concept of a 'class' or the operation of 'classification' is an internalised version of the action of grouping together objects recognized as similar. Having learned to pick out all the yellow counters in a heap and *place* them together in one spot, the child acquired the ability to *think* of all yellow objects together and thus form the concept of the 'class of all yellow objects'." The transition from perception to conception can be picked up from our analysis of the qualitative responses (Table 6). Leaving aside the edibility aspect of plants, there does seem to be a deficiency of the ability to form abstract classes below Class III. Now Class III, as has been noted, had an average age of 9½ years; and Class II's mean age was 8 years. Even allowing for the difficulty of absolute precision in age reckoning, it would seem that there is a little lag in years for this period of 'concrete operations.' That these African children reach it, whether literate or illiterate, is certainly the case. However, our figures suggest that, in comparison with European standards, already there is a difference, on the assumption that we can equate the material used with those of other workers. If valid, this is an important finding because there is no difference between literates and illiterates when not using material upon which the factor of training would be likely to influence differences, i.e., blocks and designs. Using familiar material which they have had an opportunity of manipulating, the evidence suggests that the growth in the capacity to reach the state of 'concrete operations' proceeds at a similar rate in both school and illiterate children.

VIII.—REFERENCES.

[1] Carothers, J. C. (1953) *The African Mind in Health and Disease.* Geneva World Health Organisation.

[2] Haward, L. C. R, and Roland, W. A. (1954) Some inter-cultural differences on the Draw-a-Man test: Goodenough scores *Man*, 54, 86-88

[3] Haward, L. C. R., and Roland, W. A. (1955a). Some inter-cultural differences on the Draw-a-Man test: Part II, Machover scores *Man*, 55, 27-29.

[4] Haward, L. C. R., and Roland, W. A. (1955b). Some inter-cultural differences on the Draw-a-Man test. Part III, conclusion. *Man*, 55, 40-42.

[5] Wintringer, G. (1955). Considérations sur l'intelligence du noir africain. *Revue de Psychologie des Peuples*, 10, 37-55.

[6] McConnell, J. (1954). Abstract behavior among the Tepehuan. *J. Abnorm. Soc. Psychol.*, 49, 109-110.

[7] Jahoda, G. (1956). Assessment of abstract behavior in a non-western culture *J. Abnorm. Soc. Psychol.*, 53, 237-243.

[8] Maistriaux, R. (1955). La sous-évolution des noirs d'Afrique. Sa nature, ses causes, ses remèdes. *Revue de Psychologie des Peuples*, 10, Part I, 167-189, Part II 397-456.

[9] Leeper, R. (1951). Cognitive Processes. In Stevens, S. S. (ed.), *Handbook of Experimental Psychology*. New York: John Wiley and Sons.

[10] Lévy-Bruhl, L. (1926). *How Natives Think*. English translation. London.

[11] Goldstein, K., and Scheerer, M. (1941). Abstract and concrete behavior. An experimental study with special tests. *Psychol., Monogr*, 53, No. 2.

[12] Hanfmann, E., and Kasanin, J. (1942). *Conceptual Thinking in Schizophrenia.* New York: Nervous and Mental Disease Monographs, No. 67.

[13] Humphrey, G. (1951). *Thinking*. London: Methuen and Co.

[14] Vinacke, W. E. (1951). The investigation of concept formation *Psychol. Bull.*, 48, 1-31.

[15] Abraham, R. C. (1940). *The Tiv People.* Second Edition, London: Crown Agents for the Colonies.

[16] BOHANNAN, P. (1954). *Tiv Farm and Settlement.* London · H.M.S.O., Colonial Research Studies, No. 15.

[17] ROLLINGS, P. J. (1960). A note on the cultural direction of perceptual selectivity *Proc. 16th International Congress of Psychology.*

[18] DUIJKER, H. C. J., and FRIJDA, N H. (1960). *National Character and National Stereotypes.* Confluence Surveys of Research in the Social Sciences, Vol. I Amsterdam . North-Holland Publishing Co.

[19] BERLYNE, D. E. (1957). Recent developments in Piaget's work. *Brit. J. Educ. Psychol.*, 27, 1-12.

ACKNOWLEDGEMENTS. This study was one of a series of investigations during a field research in 1959, sponsored from a Ford grant by the London School of Economics. The writer wishes to thank the educational authorities at Kaduna, Nigeria, for permission to visit schools in the Tiv division and to the Headmaster and teachers at Mbara Primary School for their help and kindness in taking up so much of their free time. The preparation of the manuscript was greatly assisted by comments on a previous draft by Dr. Gustav Jahoda, of the University of Glasgow, and on the linguistic aspects by Dr. David Arnott, of the School of Oriental and African Studies, the author's first Tiv teacher. Lastly, thanks are due to the writer's host in Tiv country, Clan-head Mode, in whose compound the research was conducted, and without whose authority and hospitality, might have been impossible.

(*Manuscript received, February 10th,* 1961.)

A PILOT EXPERIMENT IN REMEDIAL READING AT THE HOSPITAL SCHOOL, LINGFIELD, 1957-60

By LESLIE E. KEATING

I.—INTRODUCTION.

It is nowadays regarded as obvious that individual attention should be given to each member of a class of retarded children, and a pilot scheme aiming at the teaching of remedial reading with a ratio of *one teacher per child* has been developed at the Hospital School for Epileptic Children, Lingfield, Surrey.

The specific purpose of the scheme was to improve the reading ability of boys in the Senior (11 plus) Department of this Special School who were unable to read (possibly having missed a great deal of schooling) or whose reading was too far retarded to enable them to fit into a class otherwise more suited to their general social and academic standard.

II.—PHASE I.

In the first phase of the experiment, intensive teaching of reading was tried *within a very retarded class*, with a group of children who were considered most likely to progress with intensive coaching and attention. This was doomed to failure from the start, as the retardation of the bulk of the class was such that the residue were unable to work on their own whilst the selected group had individual coaching. Various projects were tried to find a way of overcoming this, but without avail. Nevertheless, it was possible to concentrate on two boys in particular, and on the one hand, to bring them up from complete inability to read to a modest competence, and on the other to derive much experience on which to base subsequent phases of the scheme.

Composition of the class:

Number of boys in room at a time: 20.
Age range: 12y. 4m. to 16y. 2m.; Mean: 14y. 8m.; Median: 14y. 10m.
I.Q. Range: 41 to 77; Estimated Mean: 56 (not all I.Qs. available).
Range of reading retardation: 5·6 years to 9·6 years.

Results.

The eventual factor in determining progress in reading has been the *Gradient* or *Rate of Improvement* obtained by *dividing the increase in reading age by the time of coaching.*

The results in Phase I are not impressive, although eight boys obtained a gradient of 1·0 or more, indicating that they have made at least 'normal' improvement as a result of intensive coaching. Five boys, on the other hand, with a gradient of less than 0·3 were thought unlikely (with the aid also of diagnostic tests of possible organic brain injury—Strauss and Lehtinen, 1947) ever to learn to read with any form of class teaching.

III.—PHASES II AND III.

The principle on which therapy was based in Phases II and III was changed from that of Group teaching within a Class to that of 'reading at mother's knee.' (*This is the crux of the present paper*). Each boy sat with the teacher for 15-20 minutes of *undivided attention* per day. For the remainder of the time he occupied himself with silent reading practice, prescribed exercises designed to help his reading, or such other activity as the teacher could devise for him to work at on his own.

Subject to this principle, the technique was completely eclectic. Considerable attention was paid to determining the book which each boy should read at each stage of his reading (Daniels and Diack, 1959) and he was then encouraged to read as much and as fluently as possible *by prompting him wherever necessary.* In some cases a script book was specially written for a particular boy (e.g., ' Dan Dare ' readers). Look-say and phonic methods were used wherever applicable, but neither these nor any other special techniques were adopted in whole or part as a ' method of teaching.'

A large supply of different books and series of books was requisitioned, especially supplementary readers, so that in addition to progressing through a sequence of graded books, each boy had *plenty of reading practice at each level.* This has been found a vital and fundamental need.

Copy-writing exercises, sentence-building with word-cards, card matching exercises, flash-card drill and comprehension exercises (starting with simple card/envelope and yes/no questions) were geared to the Reader of the moment.

A co-ordinating background was provided by a concurrent drill in a standard ' frequency list ' (Ellis, Dolch, etc.) containing, in this case, 100 words comprising 50 per cent. of all normal reading matter. (66·4 per cent. tested on a normal library book for a 9-year-old). These words were written, as learnt, on cards carried about by the boy in a box containing also cards with his own ' personal difficulty words ' for constant revision and test.

A great deal of attention was directed to emotional factors, and a very strong, almost exaggerated *esprit-de-corps* developed. Most of the boys selected for the remedial reading unit turned out to have emotional traits specifically militating against progress in reading, such as inferiority consciousness, lack of self-confidence in general, acquired boredom, lack of motivation, conversion hysteria, etc. In addition, more personal and intimate factors were operative. Therefore, as far as possible, the atmosphere of a classroom was avoided, and the children were encouraged to read *with* the teacher as they would with a parent, e.g., for a bedtime story.

IV.—PHASE II.

Composition of class :

Number of boys in room at one time : 10.

Age Range : 11y. 6m. to 14y. 11m. Mean : 13y. 5m. ; Median : 13y. 2m.

I.Q. range : 37 to 99 ; Mean : 70.9 ; Median : 70.

Range of reading retardation : 5 years to 10 years.

Results.

The essential factors of the course appeared to be :

(*a*) The initial reading standard is, *per se*, immaterial.

(*b*) The boy must be of sufficient general intelligence to take his part in the practice and planned self-tuition which intersperses the instruction periods.

(*c*) He must possess the will to work and to co-operate fully in the scheme of reading and the practices devised for him.

(*d*) He must be free from such organic brain lesions as would preclude him from mastering the visual and motor techniques required in reading.

(*e*) It is not reasonable to expect too intensive concentration on one subject reading, even if it be broken up into diverse component exercises ; ' staleness ' rapidly results.

(*f*) Instruction should be in *short* but *frequent* and especially *regular* periods.

(*g*) Children should be selected or recommended by their own class teacher (who alone knows them well enough to diagnose) on the grounds that they have the necessary ability and will power. They can then be ' screened ' by the unit therapist.

Only one boy progressed sufficiently well in Phase II to merit return to a normal stream, but eight showed sufficient promise to continue with the Unit. The remaining twelve left, were transferred to hospital, or were dropped as ' unlikely to profit ' from this particular scheme.

V.—Phase III.

Composition of the class :

Number of boys in room at a time : 5 (four groups per day).
Age range : 12y. 7m. to 15y. 4m. Mean : 13y. 8m. ; Median : 13y. 7m.
I.Q. Range : 60 to 93 ; Mean : 76·6 ; Median : 75.
Range of reading retardation : 3y. 8m. to 9y. 5m.

In September, 1960, a start was made on what seems to be the ultimate phase of the experiment. Four groups of five boys each were admitted to the Unit for one period per day of coaching on the same basic principles as in Phase II, *returning to class for the remaining three-quarters of the day for normal classwork.*

Transfer has been made from a full-sized classroom to a small cosily-furnished sitting-room. One child at a time reads with the teacher, whilst the other four work in four cubicles (as found in a polling booth) in conjunction with a wall-full of open shelving containing ancillary material which enables unsupervised yet directed activity between personal sessions. Much of this material is ' edumatic '—i.e., of a self-checking nature, often in a pre-programmed automatic ' machine.'

Results.

There seems no doubt that the time and thought expended on this project has been well-spent, and the system as evolved in Phase III has been adopted as routine in the School.

In the first term six boys were returned to a normal class with adequate reading ability to cope (R.A. 10–11 years). One boy left, and one was transferred as ' unlikely to progress.' The remaining eleven boys were all accepted for continuation the following term with a very good prognosis.

The I.Qs. in Phase III appear to suggest some correlation between higher I.Q. and improvement in reading. In fact, the coefficient of correlation between I.Q. and Gradient as found by Spearman's Ranking method over the whole experiment (n=45) was 0·44 : a definitely significant figure.

VI.—SUMMARY OF RESULTS OVER THE WHOLE SCHEME.

It is most important to realise that qualitative as well as quantitative estimates are possible, but these do not lend themselves to statistical treatment. In every case, however, the class teacher's subjective appraisal of the quality of a boy's reading has confirmed or exceeded the claim made quantitatively.

ACKNOWLEDGEMENTS: The writer wishes to express his thanks to the Medical Superintendent, the late Dr. D. G. Kennedy, for permission to publish this paper; to Mr. E. Donovan, the Headmaster, for permission and encouragement to undertake the scheme: and to those of his colleagues who recognised the possibilities of the scheme and gave unstinted co-operation and encouragement, especially Mr. R. J. Grove, sometime Head of the Senior Boys' Department.

VII.—APPENDIXES.

TABLE OF RESULTS

	PHASE I	PHASE I	PHASE III
Number in class at one time	20	10	5
Duration of Phase	2 years	1 year	4 months*
Chronological Age Range and Median	11·4 to 16·2 14·10	11·6 to 14·11 13·2	12·7 to 15·4 13·7
I.Q. Range and Median	41 to 77 56	37—99 70	60—93 75
Range of Retardation	5·6 to 9·6	5 to 10	3·8 to 9·5
Gradient Range and Median	0·06 to 3·0 0·9	0·2 to 6·3 2·4	0·3 to 7·5 3·4
Gradient Analysis:			
7 times normal	—	—	1
6 times normal	—	1	2
5 times normal	— } 1	2 } 7	3 } 12
4 times normal	—	—	—
3 times normal	1	4	6
2 times normal	1	4	1
1 times normal	6	5	1
Less than 1 times normal	13	2	4
Ineducable	7	3	—
Returned adequate	0	1	6
Continue with Unit	7	8	10
Unlikely/Unwilling	7	9	1
Left or Hospitalised	14	3	1

*—Uncompleted year in progress.

VIII.—REFERENCES.

STRAUSS, A. A., and LEHTINEN, W. E. (1947). *The Brain-injured Child.* New York: Grune and Stratton.

RICHMOND, W. K. (1961). Teaching by machine. *Times Ed. Supp.*, 198.

KEATING, L. E. (1960). Epilepsy and intelligence. *J. Ment. Sci.*, 106, 1042.

DANIELS, J. C., and DIACK, H. (1959). *The Standard Reading Test.* London: Chatto and Windus.

(*Manuscript received, April,* 1961.)

A SUMMARY OF A STUDY OF THE READING AGES OF CHILDREN WHO HAD BEEN GIVEN REMEDIAL TEACHING

By K. LOVELL, E. JOHNSON AND D. PLATTS
(*University of Leeds*)

SUMMARY. Children who had received remedial teaching in reading by the staff of a Child Guidance Centre of an English County Borough between January, 1955, and March, 1960, were traced, and their reading ages assessed. A summary of the data obtained is given, together with the interrelationships found between a number of the relevant variables.

I.—INTRODUCTION.

IN recent years many doubts have been raised concerning the effects of what is commonly called ' remedial ' teaching in basic subjects. It is not proposed to review the growing American and British literature on this problem ; suffice to say that very little is known about the long-term effects of such teaching. What, for example, is the level of attainment in reading of pupils, two, three or four years after cessation of the treatment. The groups that have been followed-up have been small, and the retesting of children has usually been six, twelve or up to twenty-four months after the special teaching ceased. For example, recent British work includes a sample of twenty-four children followed-up for fifteen months (Collins, 1961), and a group of eight children followed up for eight-and-a-half months (Pringle, 1960)

One of the authors (K.L.) was asked by the senior educational psychologist of an English County Borough to establish the reading ages of children who had received remedial teaching in reading through the services provided by the child guidance centre. The present study was carried out in response to this request.

It must be made clear at the outset that this investigation does nothing to resolve the controversy as to whether backward children who do not get special treatment eventually reach the same standards as those who do. Nor does it provide information as to whether or not children's attitudes to school work, and their general motivation and morale, improve under such teaching. Benefits from such teaching may well accrue that are not measured by the assessment of reading ages. But the study does throw light on the likely educational progress and eventual performance, during school years, of children who are backward in reading and who are currently referred for remedial teaching.

II.—SAMPLE STUDIED AND PROCEDURE.

Data relating to all children who had been referred to the Child Guidance Centre on account of reading difficulties over the period January, 1955, to March, 1960, were obtained from records at the Centre. All children had been given the Schonell Graded Word Vocabulary Test R1 at the time of referral, and of these 259 had been accepted as suitable for remedial reading *and* had received special teaching for a period of at least three months. Pupils who did not receive treatment for at least three months were not included in this study.

Remedial reading was provided by the child guidance centre in two ways. When there were a number of children in one school at any one time who were deemed to be suitable for treatment, teaching was given in small groups at the school by a remedial teacher who was attached to the Centre. On the other hand, when the number of pupils in a school who needed treatment was so small that it would be unprofitable to send a remedial teacher to the school, children were taught at the Centre individually. Pupils received one period of remedial reading per week, each period lasting 45 minutes, whether attending the Centre for individual work or whether taught in small groups at school. It is obvious that when the remedial teachers work in the schools, they are more likely to influence the other members of staff and their methods are more likely to be continued after they themselves have ceased work in the schools. The Schonell Graded Vocabulary Test R1 was also given at the end of the period of special teaching.

Of the 259 children, ninety-nine had been taught individually at the child guidance centre, and 160 in small groups in their own schools. Unfortunately, it was not possible to trace all of the children. In the case of pupils taught at the Centre, eleven no longer lived in the county borough, two were absent from school during the retesting period, seven were in residential special schools in other parts of the country, four could not be traced by the school attendance department, and one had died, leaving seventy-four who were retested. Children taught individually will be referred to as Group A. Of the pupils taught in their own schools, eight no longer resided in the county borough, six were absent from school during the follow-up period, one was in a residential special school in another part of the country, and nine could not be traced by the school attendance department, giving 136 who were followed-up. Children taught in groups will be referred to as Group B.

In Group A there were fourteen girls and in Group B fifty-two girls, although the data for the two sexes have not been analysed separately; while on retesting, four pupils from Group A and eleven from Group B were found in day E.S.N. special schools.

Intelligence quotients taken at the time the pupils were considered for remedial teaching were available. In Group A, seventy of the I.Qs. were obtained from the Terman-Merrill revision of the Binet test or WISC, and four from the Otis Quick Scoring Test (Form A). Of the 136 pupils in Group B, forty-nine I.Qs. were obtained from the Terman-Merrill revision of the Binet test, two from WISC, and the remainder from the N.F.E.R. Non-Verbal Test 5, or the Sleight Non-Verbal Test.

Although it was necessary to give the Schonell Graded Vocabulary R1 test on follow-up so as to establish the change in reading age, the Vernon Graded Word Reading Test was also given as a check on the Schonell test. All the retesting was undertaken in the schools which the children were then attending, and took place during the period December, 1960—March, 1961, inclusive. Thus, no child was retested unless remedial teaching had ceased for at least twelve months.

To establish relationships between relevant variables, Kendall's tau coefficient was employed. The data did not always permit the use of product moment correlation coefficients since some of the distributions showed marked departure from normality. For example, the reading ages of the children at the time of referral to the child guidance centre gave an asymmetrical distribution; and for this distribution, the standard deviation could not be calculated. Moreover, Kendall's tau is convenient as one can calculate the relationship between one variable and a second, holding a third constant.

III.—Results.

The data obtained have been given in four tables for the sake of clarity.

TABLE I

Data Relating to the Period of Remedial Teaching

Variable	Group A N=74	Group B N=136
I.Q.	M=92·5, σ=12·1	M=89·0, σ=12·8
R.A. (in years) at time of referral ...	M=5·9	M=6·2
C.A. (in years) at time of referral. ..	M=10·1, σ=1·1	M=9·5, σ=·9
Mean Reading Quotient $\left(\frac{RA}{CA}\right)$ . ..	58	65
R.A. (in years) at end of remedial teaching	M=8·2, σ=1·3	M=7·9, σ=1·1
C.A. (in years) at end of remedial teaching	M=11·1, σ=1·2	M=10·1, σ=0·7
Mean Reading Quotient	74	78
Mean length of period of remedial teaching (years)	1.0	0·6
Extremes (in years) in period of remedial teaching	0·25—2·8	0·25—1·4
Mean improvement in R.A. during remedial teaching (in years)	2·3	1·7
Extremes (in years) of improvement during remedial teaching	Nil—6·1	Nil—4·8

TABLE 2

DATA RELATING TO THE PERIOD SINCE REMEDIAL TEACHING CEASED.

Variable	Group A N=74	Group B N=136
R.A. (in years) at follow-up (Schonell R1)	M=9·6, σ=1·7	M=9·1, σ=1·7
R.A. (in years) at follow-up (Vernon Graded Word)	M=9·2, σ=2·1	M=8·7, σ=1·8
C.A. (in years) at follow-up	M=13·5, σ=1·2	M=12·3, σ=1·4
Reading Quotient (R.A. from Schonell R1)	71	74
Mean progress (in years) since remedial teaching ceased	1·4	1·2
Mean length of time (in years) between end of remedial teaching and follow-up	2·4	2·2
Extremes of progress made (in years) since remedial teaching ended :		
12—23 months previously	−·6 to 5·2 M=1·0 N=27	nil to 3·9 M=·9 N=58
24—35 months previously	·1 to 3·0 M=1·1 N=25	·3 to 4·5 M=1·7 N=38
36—47 months previously	·2 to 5·0 M=1·9 N=16	·2 to 4·2 M=1·6 N=30
48+ months previously	1·5 to 5·0 M=3·0 N=6	·7 to 4·6 M=2·1 N=10

TABLE 3

DISTRIBUTION OF READING AGES (SCHONELL R.1) AT FOLLOW-UP.

	Reading Age in Years										Total
	—5·4	5·5—6·4	6·5—7·4	7·5—8·4	8·5—9·4	9·5—10·4	10·5—11·4	11·5—12·4	12·5—13·4	13·5—14·4	
Group A	0	1	6	12	19	16	8	6	4	2	74
Group B	2	8	10	31	28	28	17	10	2	0	136
											210

TABLE 4

INTER-RELATIONSHIPS BETWEEN RELEVANT VARIABLES USING KENDALL'S TAU COEFFICIENT.

Relationship	Group A	Group B
Scores on Schonell R1 and Vernon Graded Word	·91**	·93**
I.Q. and progress made during remedial teaching	·35**	·26**
I.Q. and R.A. (Schonell R1) on follow-up	·43**	·25**
I.Q. and progress made since remedial teaching ceased	·15*	·12*
Length of remedial teaching and progress during remedial period.	·30**	·28**
Length of remedial teaching and progress made since remedial teaching ceased	—·05	—·01
Progress made since remedial teaching ceased and time elapsed since	·47**	·42**
R.A. at commencement of remedial teaching and progress made during that teaching.	·25**	·19**
Progress made during remedial teaching and R.A. at follow-up	·44**	·60**
I Q. and progress during remedial teaching holding constant the length of remedial teaching	·34†	·19†
I.Q. and R.A. on follow-up holding constant progress made during remedial teaching	·38†	·12†

*—Indicates significance at 5 per cent. level.
**—Indicates significance at 1 per cent. level.
†—It is not possible to test the significance of $\tau_{xy.z}$.

IV.—DISCUSSION.

The results may be conveniently discussed under a number of headings. However, it must not be forgotten that there are great individual differences in progress made during remedial teaching and since, that are masked when referring to means.

(1) The reading ages obtained on the Schonell Graded Vocabulary test and the Vernon Graded Word test correlated highly within each age group. The pupils scored somewhat higher on the former test; this might have been due in part to familiarity with the test.

(2) For Group A the improvement in reading age per unit time of remedial teaching was 2·3, for Group B 2·8. These gains are comparable with those reported in other studies. The pupils in Group A were older and rather more backward than those in the Group B at the beginning of treatment, although at the end of the remedial teaching the mean reading quotients for the two groups were roughly the same.

(3) It is gratifying to find that on follow-up, forty-seven pupils in Group A (63 per cent.) and sixty-eight of Group B (50 per cent.) had reached a reading age in Schonell R1 test of at least nine years. It is certain that more of these children will reach this reading age before leaving school. It must be remembered that on follow-up, the mean age of Group B was 1·2 years lower than the mean age of Group A.

(4) Since remedial teaching ceased, the improvement in reading age per unit time was ·58 for Group A and ·55 for Group B, and the mean reading quotients for the two groups remain almost constant between the end of remedial teaching and follow-up.

(5) There is a small positive relationship between I.Q. and progress made during remedial teaching even when the length of treatment and progress made during treatment are kept constant. Likewise, the relationship between

I.Q. and reading age on follow-up is significant, the tau value being somewhat greater in Group A than in Group B as one would expect since about half the I.Qs. in the latter group were obtained from non-verbal tests.

(6) Those who made most progress during the period of remedial teaching tend to have the highest reading ages on follow-up ($\tau = \cdot 44$ and $\cdot 60$). But there are many exceptions. Some make no progress during treatment but improve afterwards and *vice versa*.

(7) There is a significant relationship between progress made and the length of time elapsed, since remedial teaching ceased ($\tau = \cdot 47, \cdot 42$). Moreover, whereas the distribution of reading ages at the time of referral was markedly asymmetrical, and the distribution at the end of remedial teaching approximated to a normal one ($\sigma = 1 \cdot 3, 1 \cdot 1$), at follow-up the spread of reading ages had been further increased ($\sigma = 1 \cdot 7, 1 \cdot 7$). These facts suggest that maturation is playing a part in the reading improvement both during remedial teaching and during ordinary school teaching.

(8) While it is heartening to see the progress that many children made between the time of referral and follow-up, so that rather over half had reached a reading age of 9 years and may be regarded as functionally literate, backwardness in reading at the junior school stage is a serious matter. Of the children who were in junior school at the time of referral—the majority—not a single one was found in a grammar school or in the ' A ' stream of a secondary modern or all-age school at follow-up. The most successful pupil academically was found in the ' B ' stream of a secondary modern school. He was regarded by his teachers as among the ablest in his class, with some facility in French. At follow-up there were eight pupils in Group A, and six in Group B with reading ages of 12 years or more and it is certain that more will reach this reading age before leaving school. Generally, however, we may say that a few backward readers at the junior school stage will make rapid progress at some point and even become able pupils academically (cf. Fresnel the physicist, who was a poor reader at 8, and John Hunter, the surgeon, who was unable to read fluently until 16). But most backward junior school pupils will make important though limited progress, and some will make little or none either during or after remedial teaching.

(9) The study gives little help in picking out those who are most likely to improve under remedial teaching of reading. The relationship between I.Q. and progress is too small to be of value in prediction. The pupil with the highest reading age in Group B at follow-up had the lowest I.Q.

(10) Since there seems little difference between the results of individual and small group teaching, the latter approach is to be preferred since it is more economical in terms of teaching manpower, and it brings the methods of the specialist teachers to the notice of the school staff.

Finally, we acknowledge the kindness of the Chief Education Officer and the Senior Educational Psychologist of the County Borough in allowing us to publish our findings.

V.—REFERENCES.

COLLINS, J. E. (1961). *The Effects of Remedial Education*. London : Oliver and Boyd for University of Birmingham Institute of Education.

PRINGLE, M. L. K., and SUTCLIFFE, B. (1960). *Remedial Education—An Experiment*. Caldecott Community and the Department of Child Study, the University of Birmingham Institute of Education.

(*Manuscript received, May 16th*, 1961)

CRITICAL NOTE

The Incidence of Some Supposedly Adverse Family Conditions and of Left-Handedness in Schools for Maladjusted Children : by M. L. Kellmer Pringle

This *Journal*, June, 1961.

By R. R. DALE

(*Department of Education, University College of Swansea.*)

Dr. Pringle's interesting article was first delivered as a paper at the 1960 Annual Conference of the British Psychological Society. On that occasion the writer of this note ventured a suggestion which he hoped would result in some modification of a section of the findings before they appeared in print. As the hope has been in vain he feels obliged to re-iterate the suggestion, though not wishing to decry the real merits of the article.

Dr. Pringle's findings rest on a comparison of the incidence of certain adverse family conditions amongst ascertained children in special residential schools, with the incidence of these conditions among the relevant general population. The key word, a very tricky one—is ' incidence.' In this note comment will be confined to the use of this term (or similar terms) with regard to only one of the adverse family conditions, namely divorce.

The article describes the sample of ascertained children as about half of their population in special residential schools, and these are rightly taken as representative of the total resident population of such children, aged 6 to 16+. The incidence of divorce among their parents is then calculated as a proportion of the total number of children. So far so good, but when comparison is made with the incidence of divorce in the general population, the question arises whether the statistic which is used to quantify the incidence is strictly relevant. Dr. Pringle describes this statistic as the 'divorce rate per annum' and states that this, at about 1·2 per cent., is only one-thirteenth as high as the incidence among the ascertained children. One wonders what exactly this 'divorce rate' really is, and it is quite conceivable that a detailed explanation by Dr. Pringle would clear things up. The numerator of the fraction appears to be the number of divorces per annum, but what is the denominator ? If the latter is the total population or total of married people this ' divorce rate' is inapplicable because the proportion of divorced people in the population would increase each year up to a certain limit, and would be far greater than this ' divorce rate.' If, on the other hand, the denominator is, more relevantly, the number of marriages per annum, it appears that the resulting percentage should be considerably greater than 1·2.

A few figures are given by way of illustration. The Report of the Royal Commission on Marriage and Divorce (1956) gives the crude divorce rate per 1,000 total population as ·69 for 1953, and a roughly similar figure for previous years, but as explained above, this statistic is not applicable. The number of divorces per 1,000 married population (1·9 in 1958, according to the Registrar General's Statistical Review) appears also inapplicable. A more relevant figure is the proportion of marriages terminated by divorce, where the marriage denominator is the average number of marriages contracted annually five to fifteen years earlier. For 1954 this figure expressed as a percentage was 6·7. It will be noted, however, that even a re-adjustment of the figures on some such scale would not invalidate Dr. Pringle's main conclusions.

There are a few incidental points which Dr. Pringle might like to discuss in her reply. Might not foster-parents and step parents be much more willing to have a troublesome child ascertained than would the real parents ; and might not such children, once sent away to a residential school, be kept there much longer than children who had more normal homes to which to return ? Finally, perhaps Dr. Pringle would find additional interesting results if she analysed the social class composition of her ascertained children and then compared the incidence of adopttion, illegitimacy, etc., with that in a control group.

A Reply to R. R. Dale's "Critical Note."

BY M. L. KELLMER PRINGLE

I would like to thank Mr. Dale for giving me the opportunity to reiterate in print (and thus in greater detail) the answer I gave to him verbally during the 1960 Annual Conference of the British Psychologocial Society. Perhaps I ought to have included it in the original article rather than merely stating as I did, that "precise information is lacking on the incidence of broken homes amongst the population in general" (p. 187).

The only really satisfactory way of arriving at a reliable estimate would have been as follows :

(*a*) To find out the year in which the parents of each of the maladjusted children in my study had been divorced.

(*b*) To find out the ages of the children (if any) of the parents in the general population divorced during the years in question in order to calculate for each of those years the divorce rate of parents with children in the relevant age groups

(*c*) Then a mean percentage or range for the general population over the ten-year period could have been compared with the figures found in my sample of maladjusted children. However, the information needed under (*a*) would have been almost impossible to obtain without a visit or at least a letter to the parents of all the children concerned ; personal visits were quite impracticable and it is highly doubtful whether many parents would have replied to a postal enquiry.

A second-best estimate could have been obtained by comparing the mean figure worked out for the population in general as in (*c*) above and comparing this simply with the mean for the whole maladjusted group of children. After consulting all official sources, including information published by the United Nations, I came to the conclusion that the figures I needed were not readily available. Therefore, I wrote to the Registrar-General asking for this information or alternatively for access to data from which I could myself extract it. I hoped to obtain comparable data not only for divorce but all the other adverse family conditions, the incidence of which I was investigating in schools for maladjusted children. However, neither request was conceded. When mentioning this at the Annual Meeting in 1960, I coupled it with a plea that any such information, which must be available as a result of various Government surveys, should be made available to *bona fide* research workers.

The method suggested by Mr. Dale for obtaining comparative data for the normal population, would be beset by even greater difficulties since I would need to know (*a*) the year when the parents of each of the maladjusted children were married and (*b*) how many of the people in the general population, who were married in the same years, subsequently had children of the same ages as those in my sample of maladjusted children. I do not know the former and the latter would again require official statistics which are at present not being made available. Though Mr. Dale's method of calculating the divorce rate would be rather more complicated than the way I had intended to make comparisons, I am not certain that it would be more reliable than mine. However, since neither method is at present practicable, this is a theoretical issue.

Thus, circumstances forced me to adopt what can only be considered a third-best method : this was to take the mean divorce rate per 1,000 married population for the period 1943 to 1958 and then reduce this figure since it included childless couples as well as parents with children aged under 6 and over 16 years (ages not included in my maladjusted sample.) The size of this reduction could only be guesswork and for this reason I quoted no exact figure (such as Mr. Dale claims I did) but indicated how much of an estimate it was by giving for the general population incidence "1 per cent. to 1½ per cent." in Table 4 (p. 191). I was emboldened to use such a rough estimate and quote such an imprecise figure by the fact, which Mr.

Dale is kind enough to mention himself, that, however I tried to readjust the figures (such as only looking at the excessively high divorce figures just following the war), the incidence of maladjusted children whose parents' marriage had ended in divorce was still very much higher than in the general population.

Now to turn to the points raised by Mr. Dale in the last paragraph of his note: I think it is quite likely that adoptive parents (I did not deal with foster-parents) and step-parents are probably "more willing to have a troublesome child ascertained than would the real parents"; moreover, there is also an increased likelihood that a higher incidence of emotional disturbance exists among adopted children because of some inherited instability or as the result of early deprivation or institutionalisation. I briefly touched on this point when considering the concept of the broken home (p. 188), but lack of space prevented a fuller discussion. Secondly, while it is possible that children who have adopted or step-parents are kept longer in residential schools than children who have more normal homes, I do not think that this is necessarily so. The shortage of places in residential schools for the maladjusted and the pressure of long waiting lists is such that in most cases known to me (and their number is, of course, limited) children are returned to their homes as soon as they have improved sufficiently and their parents have modified or changed their attitude to the child sufficiently to make a recurrence of severe problem behaviour seem unlikely. With regard to Mr. Dale's final point, I am afraid I do not have any information on the social class composition of the ascertained sample, but quite agree that a comparison of the incidence of adoption, illegitimacy, etc., with that in a control group would be of interest.

RESEARCH NOTE

Note on the Problem of Inference in Educational Research.

By C. D. HARDIE
(University of Tasmania)

One of the most striking characteristics of much of the research that takes place in modern education is its dependence on numerical data. Now there are always two sorts of question which can be asked about such data. The first is concerned with how far things (or qualities or states) to which the data refer are capable of measurement; the second asks what inferences may be drawn from the data. I do not believe that nearly enough attention has been devoted to either of these questions, but this note is confined to the second.

The inferences which are usually drawn from such data are also of two sorts, according as they confine themselves to or go beyond the data. Inferences drawn about a class of school children from sets of examination results may be of the former kind, and do not raise any particular problems. But when the inferences drawn go beyond the data, then the question of their justification does arise. The inferences go beyond the data in all cases where a conclusion is reached about a population which is wider than that to which the data refer. Such a conclusion is put forward first of all as a hypothesis (often called the null hypothesis), and an attempt is made to see if the data render the hypothesis unlikely.

A typical instance, thoroughly familiar to readers of this *Journal*, is where the data consist of two sets of measurement, and the inference is drawn that a significant difference exists between the means of the two sets. What is the meaning of such a statement? The official answer roughly is that a difference is significant when it is not due to chance. More precisely, it is that if the two sets of measurements had been measurements of random samples from the same population (the null hypothesis) then the probability of getting as big a difference as we have actually got is less than 0·05, or 0·01, or whatever ' level of significance ' is chosen. At first sight this may appear clear enough, but difficulties arise immediately we try to analyse the meaning farther. The two key ideas are ' random sample ' and ' probability,' and a little consideration soon raises doubt about the validity, or even the meaning of many of the inferences that have been drawn in educational research

RANDOM SAMPLE.

A random sample of any population is generally said to be one in which every member of the population has an equal chance of being drawn. In the light of such a definition, consider the famous investigation of the Scottish Council for Research in Education into the intelligence of Scottish children. In 1932 and 1947 the same group test was given to 11-year-old Scottish children. The Council published the results in Table X of *The Trend of Scottish Intelligence* and comment " the differences between the 1947 and 1932 results, and between the 1947 boys and girls, as shown in Table X, are all statistically significant, and the probability of their being chance differences is almost infinitesimal."* What does this mean? In view of the above explanation of significance, it ought to mean that the 1932 and 1947 results cannot be regarded as random samples from the same population. But if we ask from what population, there is no obvious answer. Moreover, even if some hypothetical population were imagined, such as the population of 11-year-olds in Scotland throughout all time, it is clear that the 11-year-olds of 1932 and of 1947 are not random samples in the sense explained. We must always remember that when significance has been established and we have rejected the idea that a very unlikely event has occurred, then there are two possibilities open: (1) that the sets of measurements do not come from the same population; (2) that they are not random samples. We cannot be forced to accept (1) unless we have excluded (2). But in the investigation into the intelligence of Scottish children, we must accept (2). Consequently, the application of statistical weapons to the data in this case seems to me to have no clear meaning.

---

* The Scottish Council for Research in Education, *The Trend of Scottish Intelligence* (University of London Press, 1949), p. 87.

PROBABILITY.

There is a long history of controversy about the meaning of probability, but for the present discussion we can confine the issue to two views. The first is that associated with Johnson and Keynes, and represented now mainly by Jeffreys. The second is that associated with Karl Pearson and Yule, and represented now mainly by Fisher. E. S. Pearson and Neyman have diverged considerably from this second view, but they are still much closer to Fisher than to Jeffreys.

For Jeffreys, probability is a logical relation between two propositions. We can deduce the probability of a sample from a population of given character, but when we wish to infer the probability that the population has a certain character from a sample, then we need to use inverse probability as provided by the theorem of Bayes. This states that when we have a number of hypotheses about the population which are mutually exclusive, and such that one of them must be true, then the probability that a particular hypothesis is true after the sample has been observed can be expressed as a function of the probability of deriving the sample on the various hypotheses and of the prior probabilities of the hypotheses.

In educational research, no attention, as far as I know, has been given to the views of Jeffreys. In *The Trend of Scottish Intelligence*, for example, no mention is made of the prior probability of the hypothesis that intelligence in the population is declining (or is not declining), and it is clear that such a question was not even considered by any of the contributors to that volume.

This means that in considering the meaning that is to be attached to statements of probability in educational writing, it is the view of Fisher that is taken for granted. Now no one can claim that Fisher has explained at all clearly, or without mythological trappings such as his hypothetical infinite population, what he does mean, but it can be taken as definite that some frequency interpretation has to be given to the idea of probability. Thus, when we say that the probability of getting as big a difference as we have actually got by drawing random samples from the same population is only 0·05, this must be interpreted as meaning that in a long series of experiments we would be wrong only once in twenty times if we acted in this way. But such an interpretation again makes nonsense of the application of a significance test to the comparison of the 1932 and 1947 populations.

I do not wish to belabour the authors of *The Trend of Scottish Intelligence* unfairly or to give the impression they are more illogical than some other experimenters in education. I chose that as an example because some, at any rate, of the contributors had a high reputation and because it has been considered an important publication. But the same kind of muddle is very common, as casual observation of books and journals will quickly show.

This means that more care should be taken by experimenters in education in planning their experiments, and they should not automatically apply statistical tests to the data they obtain. Indeed, such application may well be opposed to the way in which progress has, in general, been made in the sicences since the time of Bacon. Suppose that most children learn something taught by method A better than if taught by method B, while a minority learn better by method B, than by method A. If random samples of children taught by the two methods are obtained, then a significant difference will probably be found in favour of the sample taught by method A. So far so good, but what often happens then is that text-books and lecturers pronounce that method B should be superseded by method A in the schools, and the matter rests there until someone suggests method C. But where, as in the conditions under which learning takes place, so many factors are involved, it is surely naive to leave the matter there. What is required is to find those factors in the situation that are associated with better learning under method A, and those factors that are associated with better learning under method B. In this way further research is stimulated and more useful generalizations obtained.

In conclusion, we can say that when a statistical test is applied it should always be made clear (*a*) what the population is from which a sample (or samples) is assumed to have been drawn and about which a conclusion is to be reached ; and (*b*) that the sample (or samples) has been drawn at random. When that has been done, if a frequency interpretation is given to probability and a decision made about the number of times the experimenter is prepared to draw a false conclusion, the test will then decide if the null hypothesis about the population should be rejected.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESIS

A Factorial Study in the Domain of Creative Thinking

By E. E. SULTAN

(*Abstract of thesis submitted for the degree of M.A., University of London, Institute of Education*)

The recent work of J. P. Guilford, in California, on high-grade abilities indicates that numerous intellectual factors beyond general intelligence and the common group factors can be isolated and measured. In particular, he distinguishes 'divergent' or productive thinking from 'convergent' thinking to a predetermined solution, and describes tests of such abilities which can be scored with reasonable objectivity. The main factors in the general domain of creative thinking are named as:

A.—Flexibility of Closure.
B.—Speed of Closure.
C.—Adaptive Flexibility.
D.—Spontaneous Flexibility.
E.—Originality.
F.—Ideational Fluency.

The present research attempts to show whether these factors can be isolated among English 13-14 year grammar school pupils, or how far they resolve into more familiar factors such as *g*, verbal and spatial abilities. Tests were selected, adapted or constructed to measure Guilford's factors plus a number of reference factors, and administered by the writer to a total of ninety boys and eighty girls.

A.—*Flexibility of Closure*: holding in mind a configuration against distractors

1.—Gottschaldt figures—Thurstone's multiple-choice version.

2.—Copying designs on dotted spaces.

3.—Thurstone's Sigma test—marking figures in which the letter Σ is concealed.

B.—*Speed of Closure*: quickness of restructuring relatively unstructured stimuli.

4.—Thurstone's Mutilated Words—recognizing words from incomplete letters.

5.—Street Gestalt Completion.

6.—Four-letter words—rows of evenly-spaced capital letters, picking out sets of four that make words.

C.—*Adaptive Flexibility*: discovering new approaches to the solution of problems

7.—Matchstick problems—numbers of different solutions to problems involving figures made of matchsticks.

8.—Guilford's Squares test—placing X's on a checkerboard with no two in the same row, column or diagonal; number of different solutions.

9.—Ingenuity problems—reasoning problems requiring restructuring rather than routine method of problem solving.

D.—*Spontaneous Flexibility*: freedom from inertia of thinking so as to give a diversity of ideas.

10.—Brick uses (flexibility)—number of changes of category in listing possible uses for a brick.

11.—Unusual uses for certain common objects—number of acceptable alternatives.

12.—Object naming (flexibility)—writing names of 'fluids' and 'plants,' numbers of changes of category.

E.—*Originality*: production of unusual and clever responses

13.—Consequences (clever)—number of unusual and clever predictions of consequences of unusual situations.

14.—Impossibilities—number of reasonable suggestions.

15.—Anagrams (clever), made up from the letters GENERATIONS—proportion of rare solutions.

16.—Inkblots (original)—Rorschach cards Nos. 1, 3, 8, 10 shown on slides—numbers of unusual responses.

17.—English creativity. Essay topic, " The Hut in the Wood," marked by pairs of teachers for quantity and range of good ideas, originality and imagination.

18.—Art grades—school marks.

19.—Interests—questionnaire on leisure occupations, proportion of responses involving creative activity.

F.—*Ideational Fluency*: speed of calling up ideas, irrespective of quality.

20.—Brick uses (fluency)—scored for total reasonable suggestions.

21.—Consequences (fluency)—number of suggestions.

22.—Topics—numbers of ideas associated with given topics.

23.—Inkblots—total responses.

G.—*Rigidity-Flexibility*: miscellaneous tests suggested by other investigators, Luchins, Lovell, Chown, etc.

24.—Hidden words—in first list of scrambled letters, a set is established to discover names of animals; a second list also contains animal names together with other shorter and easier hidden words; score—number of the latter discovered.

25.—Mathematical approximations—a series of arithmetical, algebraic and geometric problems which can be solved either by routine methods requiring rough work written in the margin, or by quicker indirect methods; score—number solved without rough work.

26.—Concept formation—Lovell's test based on Vinacke's figures, to be classified according to two or more mutually exclusive principles.

27.—Questionnaire (Chown)—number of preferences for flexible as against rigid or obsessional activities.

28.—Double-meaning words—writing a third word which means the same as a word on the left and, in a different sense, the same as another word on the right.

H.—*Word Fluency*: rapid production of words to fulfil structural requirements, regardless of meaning.

29.—Anagrams (fluency)—number of different words.

30—Prefixes—listing words beginning with ' Con—.'

31.—First and last letters—listing words with specified first and last letters.

I.—*Perceptual Speed*: facility in perceiving detail imbedded in irrelevant material.

32.—P test—a standard figure followed by 8 figures, 3 to 5 of which are identical with the standard; number correctly underlined.

33.—Scattered x's (Thurstone)—letters scattered all over the page, picking out 7 x's on each page.

34.—Minnesota Clerical Test 2—pairs of names, to be checked as same or different.

J.—' *g* ' probably present in many of the above tests, especially C and D. Additional reference tests:

35.—A version of the Shipley Abstraction test.

36.—Non-verbal classification (Lovell)—sets of six figures, three to be underlined as alike in each set, also stating the principle of classification.

K.—' *V* ' *factor*, probably present in many of the D, F, H tests and No. 28.

37.—Vocabulary—multiple-choice tests (Chown).

L.—*Number factor*, in No. 25 and: 38—School mathematics examination marks.

M.—*Spatial or Visualization factor*, possbily in A and C.

39.—Paper formboard—creative-response version of Minnesota test.

40.—Memory for designs (Lovell's version).

ANALYSIS AND RESULTS.

Intercorrelations were computed by the London University Computer Unit, and eight principal component factors were extracted. Rotations were carried out graphically, and an independent Varimax solution of seven factors was kindly supplied by Dr. H. F. Kaiser of the University of Illinois (Table 1). These two solutions were essentially very similar.

No general factor was found running through verbal and non-verbal tests, nor a broad verbal factor in all tests involving the use of words. The probable reason for this is that grammar school pupils are highly selected for verbal intelligence and attainment. Hence, by far the most prominent factor (No. 1) is general to the

TABLE 1

VARIMAX SOLUTION.

Factor No.	I	II	III	IV	V	VI	VII	
Identification	Spa. +g	Id. Fl.	*V*	Word Fl.	Orig.	Per. Sp.	Resid.	h^2
1. Gottschaldt figures	61	04	−12	05	33	25	−17	589
2. Copying designs	76	09	−02	−02	06	12	−12	622
3. Sigma test	60	−04	−01	07	22	19	−26	515
4. Mutilated words	26	04	53	15	−04	23	−28	496
5. Gestalt completion	11	10	37	−15	41	09	−26	427
6. Four-letter words	23	−05	23	12	15	54	−08	441
7. Matchstick problems	53	−04	−03	13	10	−01	06	311
8. Squares test	41	−05	−33	22	24	−01	08	389
9. Ingenuity problems	26	−10	33	63	03	−08	−15	610
10. Brick uses (flexibility)	22	32	02	12	12	−41	03	353
11. Unusual uses	16	47	11	04	42	−13	00	452
12. Object naming	−18	18	09	40	36	−03	−36	488
13. Consequences (clever)	08	59	07	12	00	−10	10	396
14. Impossibilities	−06	14	−16	07	46	04	01	262
15. Anagrams (clever)	11	−08	−00	05	17	22	62	485
16. Inkblots (orig.)	06	10	−23	09	50	−01	23	377
17. English creativity	−10	12	−07	−06	06	−10	53	331
18. Art grades	21	−11	11	−05	47	−09	00	305
19. Creative interests	05	−03	20	−33	67	−02	17	626
20. Brick uses (fluency)	−06	48	−07	18	27	21	−25	456
21. Consequences (fluency)	04	73	12	−04	−08	08	−02	563
22. Topics	11	65	−27	04	17	10	11	564
23. Inkblots (total)	−01	23	−57	−03	26	10	10	474
24. Hidden words	09	16	05	62	−17	10	−05	456
25. Math. Approximations	56	−04	02	14	−08	−10	03	350
26. Concept formation	52	08	09	04	−09	08	−01	296
27. Flexibility Quest	−03	39	05	−37	01	01	−05	301
28. Double Meaning Words	20	09	58	03	15	−01	−02	497
29. Anagrams (fluency)	02	18	−19	28	13	47	−13	419
30. Prefixes	−05	33	02	51	−03	28	−08	450
31. First and last letters	26	−07	16	58	11	05	29	527
32. P test	35	−11	02	−18	03	45	−13	389
33. Scattered x's	25	03	42	05	−07	34	11	378
34. Minnesota Clerical	02	19	−02	13	−11	72	08	598
35. Abstraction	61	12	21	10	−10	28	22	577
36. Non-verbal Classification	51	29	23	03	−04	06	10	411
37. Vocabulary	−14	28	51	09	18	−02	17	425
38. Maths. grades	61	15	19	−10	−09	−11	21	500
39. Paper formboard	67	12	17	08	09	01	−29	602
40. Memory for Designs	66	−17	05	−13	27	02	13	566

All decimal points omitted. Loadings of ·30 and over in heavy type.

spatial, flexibility of closure, *g* and Mathematics tests. This factor is also the main one in the non-verbal Adaptive Flexibility tests, 7 and 8. No rotations were able to separate Flexibility of Closure as a distinct factor. However in the graphical solution, a small Inductive Reasoning+Mathematics factor (Tests 35, 37, 38) was separated from the spatial+closure factor.

Factor III clearly corresponds to Thurstone's *V*, though it is small and shows some discrepant loadings(e.g., Tests 5, 23, 33). In our selected group the most prominent verbal factor is No. II, clearly identifiable as Ideational Fluency, though also loading Spontaneous Flexibility (Nos. 10, 11) and the Consequences Originality Test (No. 13).

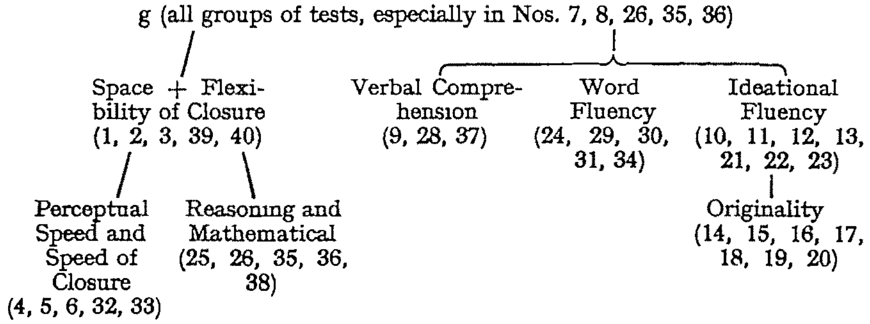
Factor IV may be identified with Thurstone's Word Fluency. It was more clear-cut in the graphical solution, showing high loadings for Tests 6, 24, 29, 30, 31 and 34 only.

Factor VI corresponds to Thurstone's P, perceptual speed ; the loadings for Tests 6 and 29 fit in reasonably with this interpretation.

So far no support has appeared for Guilford's three Flexibility factors. But No. 5 can reasonably be called an Originality factor, though it is mainly present in Creative interests (19), Inkblots original (16) and Art grades (18), rather than in the tests proposed by Guilford. However, Impossibilities (14) and two of the Spontaneous Flexibility tests (11, 12), and Gestalt completion (5) give significant loadings. In the graphical solution, Anagrams (15) fell in the same group, but here it combines with English creativity in the residual Factor VII.

GROUP FACTOR ANALYSIS.

On the basis of the sign-pattern of the unrotated factors, and the results of the graphical rotations, the correlation matrix was re-arranged to group most of the variables into seven clusters, and Burt's method for non-overlapping group factors was applied. This technique yielded a considerable *g*-factor, particularly in the Flexibility of Closure, Spontaneous and Adaptive Flexibility, as well as the expected reference, tests. However, some overlapping, either positive or negative, was clearly still present among the group factors. The most accurate picture of the structure of the battery would seem to be as follows :


g (all groups of tests, especially in Nos. 7, 8, 26, 35, 36)
Space + Flexibility of Closure (1, 2, 3, 39, 40)
Verbal Comprehension (9, 28, 37)
Word Fluency (24, 29, 30, 31, 34)
Ideational Fluency (10, 11, 12, 13, 21, 22, 23)
Perceptual Speed and Speed of Closure (4, 5, 6, 32, 33)
Reasoning and Mathematical (25, 26, 35, 36, 38)
Originality (14, 15, 16, 17, 18, 19, 20)

MAIN CONCLUSIONS.

1.—Guilford's creativity factors—Spontaneous and Adaptive Flexibility and Originality—are not, as readily measurable among 14-year-old grammar school pupils as among high-grade adults. They tend to resolve into familiar factors such as *g*, *v*, *k* and fluency.

2.—However, an Originality factor can be distinguished, showing itself more in creative leisure-time interests, Rorschach responses and art work than in written tests or subjectively marked essays. No evidence is yet available as to how predictive this is of future creative achievement.

3.—Flexibility of Closure, at least at this level, is identical with the *k* or Visualization factor, and Speed of Closure with the Perceptual Speed factor.

REFERENCES.

CHOWN, S. M. (1961). The effect of age on the relationship between different types of rigidity. *Quart. Bull. Brit. Psychol. Soc.*, No. 44, A12.

GUILFORD, J. P. (1959). Three faces of intellect. *Amer. Psychologist*, 14, 469-479.

LOVELL, K. (1955). Intellectural deterioration in adolescents and young adults. *Brit. J. Psychol.*, 46, 199-210.

SULTAN, E. E. (1961). M.A. Thesis, University of London Library.

THURSTONE, L. L. (1944). *A Factorial Study of Perception.* University of Chicago Press.

Forgetting in Emotionally Disturbed Secondary School Children

BY W. E. P. WINDER

(*Summary of thesis presented for the degree of M.A. in Psychology of the University of Liverpool, in* 1960.)

INTRODUCTION.

It is a common experience that emotional disturbance inhibits the process of learning. The discomfort of the frustration caused in this way is something with which we are all familiar; but the loss of what has already been learned, by the interference of emotional disturbance, over a period of time should be of great concern to the professional teacher, were such to be proved. Empirical evidence, that forgetting is accelerated more or less by the interference of emotional disturbance, appeared plentifully to the author and his colleagues in secondary schools It was felt that if this tendency for emotionally disturbed children to lose ordinary learned material over a given period of time could be measured, teachers might be able to make sympathetic allowances in such cases, and perhaps alleviate the conditions in making them known to the parents.

THE INVESTIGATION.

Two measurements were required; first, an assessment of the emotional adjustment of a number of Subjects, and secondly, a measure of each Subject's forgetting, using ordinary school lesson material in Word, Number, Picture and Music.

Following a number of pilot investigations, for which meaningful Word, Number, Picture and Music tests were devised, seventy-nine boys and girls in one North country secondary school and twenty-four in a Southern county school, all of 12 years of age, were submitted to testing. In the testing of the first group, however, difficulties in oral reproduction of music arose, and the results of the Music test were not utilized.

The pupils' own Form teachers, in the course of the ordinary curriculum, gave them a set prose passage to learn, measured the amount learned by immediate recall, and after an interval of twenty days again requested the recall of the passage Credit was given on each occasion for each correctly positioned word in the reproduction. A similar device was employed in Number, where a group of meaningful numerical expressions was used; in Picture, where an interesting composition of a large number of items could be readily identified; and (in the second test only) in Music, where a meaningful phrase was mingled with a number of increasingly similar phrases, and identified from the groups by recognition, this method replacing the faulty reproductions experienced in the first test. The marks obtained from the correct reproduction and recognition in these tests were totalled for each Subject after initial learning, and again following reproduction after an interval of twenty days, this being the period of time arrived at from the pilot testing. The difference obtained between the scores in the two recalls was called the Forgetting Score. It was assumed that any loss of learned material would have been caused to some extent during this period of time by, among other things, the natural play of emotional disturbances, and this tendency was measured by using the Bristol Social Adjustment Guides (D. H. Stott, University of London Press, 1958). The Guide consists of a questionnaire which is registered by the Subject's own teacher, and indicates the way in which the Subject habitually responds to his environments, showing the extent of predisposition to maladjusted responses, for future reference. The Guide results are arranged to divide into two main groups of ' unsettledness ' and ' maladjustedness ' and numerical scores obtained can be totalled to give an overall score for each Subject.

For the purposes of this research, the sections of the Guide which gave a numerical assessment were Hostility, Anxiety (toward other children and toward adults), Depression, Restlessness, Withdrawal, Abandon and Unconcern and Miscellaneous General and Undifferentiated symptoms, all graded in order of severity. One unit of score was allowed for each registration on the unsettledness side of the Guide, and two units on the maladjustedness side, irrespective of the position of the registration. This modification in scoring was found to be sufficient to provide a figure which could then be compared with the Forgetting Score, obtained from the tests. It was interesting to discover that the Social Adjustment Guide scorings in most cases compared very well with the empirical records which had been kept by the schools. This was borne out in detailed case studies on a random selection of Subjects.

Results.

The scores obtained from the Guides were referred to as Social Maladjustment Scores and were correlated with the Forgetting Scores and I.Qs. :

Social Maladjustment Score = 1.
The Forgetting Score = 2.
Intelligence Quotient = 3.

$r_{12} = +0{\cdot}42$ $r_{13} = -0{\cdot}25$ and $r_{23} = +0{\cdot}18$ in Test 1.
and $r_{12} = +0{\cdot}50$ $r_{13} = -0{\cdot}18$ and $r_{23} = +0{\cdot}04$ in Test 2.
Partial Correlations of the above results then gave :
$r_{12\cdot3} = +0{\cdot}49$ $r_{13\cdot2} = -0{\cdot}36$ and $r_{23\cdot1} = +0{\cdot}32$ in Test 1.
and $r_{12\cdot3} = +0{\cdot}52$ $r_{13\cdot2} = -0{\cdot}23$ and $r_{23\cdot1} = +0{\cdot}15$ in Test 2.

The partial correlation $r_{12\cdot3}$ in Test 1 was highly significant, and the same correlation in Test 2 was estimated to be just acceptable because of the smaller number of Subjects tested.

Inter-test correlations showed a small positive relationship of $+0{\cdot}236$ between Picture and Number test results, and a still smaller relationship of $+0{\cdot}185$ between Number and Word, which is as easily understood as the relatively large correlation between Picture and Word, of $+0{\cdot}45$.

The chief limiting conditions in the work were that it was :

(1) Applied to 12-year-old children only.
(2) Influenced by what appeared to be ' reminiscence.'
(3) Executed using the results of only three school subjects in the first case, and four in the second (i.e , with the addition of Music).

Application of the Resuls.

The relatively high positive correlation between Social Maladjustment and Forgetting suggests that tendencies to emotional disturbance make ' retention ' difficult by interfering with memory.

In cases where the Social Adustjment Guide reveals a marked level of maladjustment remedial action might be taken after discussion with the parents and, if necessary, referral to child guidance services. But the same information is clearly useful to the schools in helping teachers to differentiate between slow learners and those whose retention is affected by emotional factors ; and it would certainly be of value to allocation panels

A Comparative Study of Three Methods of Investigating the Interests and Out-of-School Activities of Children of Secondary School Age.

BY L. M. SMALLEY

(*Abstract of thesis submitted for the degree of Master in Education in the University of Manchester, September,* 1960.)

OBJECT.

The object of the investigation was to compare three existing methods of measuring the out-of-school interests of children of secondary school age.

METHOD.

The investigation was carried out during 1958 and 1959 in a boys' secondary modern school in a mixed rural-urban area of S.E Lancashire, the subjects being 125 boys aged 11 to 14 years. Their out-of-school interests were assessed by means of the three tests listed below.

(i) A diary technique, previously described by the author (Smalley, 1958) in which the boys recorded their leisure-time activities for five days on diary forms marked off in quarter hours.

(ii) A modified version of the Word Recognition Test used by Peel (1948) and Pennington (1950), in which words were divided into four groups, ' intellectual ' (I), ' passive ' (P), ' physical activity ' (A) and ' hobby ' words (H), the two latter groups being obtained by the division of Pennington's single group of ' activity ' words. Each sub-test contained six words, three from each of two of these groups and the boys were asked to identify any three of these six words The six possible combinations of groups (I v P, I v A, I vH, P v A, P v H and H v A) each occurred four times in the complete test, making twenty-four sub-tests in all.

(iii) Part of the Interest Finder used in Springfield, Missouri, by Jersild and Tasch (1949), in which the boys were asked to respond to the four items, " My three wishes," " What I like best outside school," " What I like least outside school," and " One of the happiest days of my life." Four or five blank lines were left below each item for the response and a score was entered in the appropriate sub-category (see below) for each interest revealed.

CLASSIFICATION.

Interests which could be measured on all three tests were classified into the following cagteories :

Main Categories.	*Sub-Categories.*
Physical activities.	Sports and Games, Swimming, Cycling, other outdoor activities.
Passive occupations.	Television, Radio, Cinema, Comics, Spectator Sports.
Active hobbies.	Fishing, Constructional, Gardening, Farming, Keeping pets.
Intellectual pastimes.	Reading, Quiet games, Music, Art.

RESULTS.

In each category three scores were recorded for each boy :

(i) The number of hours spent on the activity in one week (D).

(ii) The number correctly identified of the words used to represent the activity (W).

(iii) The number of wishes, likes, dislikes and happiest days recorded in that particular category (I).

These scores were used to calculate the inter-test correlations.

(*a*) *Grouped Scores*. The mean scores in the four main categories are shown in Table 1.

TABLE 1

MEAN SCORES IN THE MAIN CATEGORIES.

Category	Mean Scores			Mean score as a percentage of the total mean score		
	D	W	I	D	W	I
Physical activities	13·0	13·8	13·9	33	25	51
Passive occupations	17·5	14·5	2·9	44	26·5	11
Active hobbies	4·9	15·6	7·8	13	28·5	29
Intellectual pastimes	3·85	11·1	2·5	10	20	9

On all three tests physical activities obtained a high score, whilst intellectual pastimes recorded the lowest score. Passive occupations obtained the highest score on the Diary, the second highest score on the Word Recognition Test, but a low score on the Interest Finder. More than half the latter represented interest in watching football, none of which was recorded in the other two tests. The Diary score for active hobbies was low compared with that obtained on the Word Recognition Test and Interest Finder.

(*b*) *Inter-Test Correlations*. Where only one distribution was approximately normal, the point biserial correlation was calculated, the product moment correlation being calculated where both distributions were approximately normal. The resultant correlations are shown in Table 2.

TABLE 2

INTER-TEST CORRELATIONS.

(r) = Product moment correlation.
(p) = Point biserial correlation.

D = Diary.
W = Word Recognition Test.
I = Interest Finder.

Correlations exceeding ·230, marked ^{xx}, are significant at the 1 per cent. level; those exceeding ·174, marked ^{x}, are significant at the 5 per cent. level. (Guilford, 1956, pp. 303 and 538–9).

	W *v* D	W *v* I	D *v* I
Physical Activities	— ·178 (r)^{x}	— ·091 (p)	·007 (p)
Passive Occupations	·235 (r)^{xx}	·231 (p)^{xx}	·060 (p)
Active Hobbies.	·342 (p)^{xx}	·375 (p)^{xx}	·421 (p)^{xx}
Intellectual Pastimes	·057 (p)	·057 (p)	·187 (p)^{x}

The low and negative correlations appeared to be due to five main causes :

(i) The low reliabilities of the tests (no figures are available for the Interest Finder ; Smalley (1958) quotes reliabilities of 0·4 to 0·6 for the Diary, and Pennington (1950) reliabilities of 0·4 to 0·7 for the Word Recognition Test).

(ii) The time lapse between the tests (seven months between the Diary and the Word Recognition Test).

(iii) Inconsistencies in the children.

(iv) Differences between what the tests actually measured.

(v) Deficiencies in the tests themselves.

TEST DEFICIENCIES.

(i) The Diaries were not recorded for a sufficiently long period of time to provide a true picture of the leisure-time activities of individual boys ; five week-days and two week-ends would have been more satisfactory.

(ii) The Word Recognition Test was weakened by the use of homonyms, words associated with subjects of the school curriculum, words as likely to indicate interest in watching games as in playing them and names associated with ' out-of-date ' television programmes Too few words (fewer than four) were used to represent certain sub-categories

(iii) The number of possible responses to the Interest Finder was inadequate.

CONCLUSIONS CONCERNING INTERESTS.

(i) In all tests great interest was shown in physical activities. The tests did not succeed, however, in distinguishing between ' time killing ' physical activity, interests in spectator sport and active interest in sport.

(ii) In the Diary and Word Recognition Test, most boys manifested great interest in passive occupations, particularly in television, but few expressed this interest in the Interest Finder.

(iii) About half the boys revealed an interest in active hobbies in at least one test. In this category all correlations were positive and many were significant.

(iv) The score for intellectual occupations of any kind was very low on all three tests.

(v) The range of interests revealed by the tests was limited. As interests are largely learned, one function of the school might be to provide an environment designed to encourage a wider range of interests.

REFERENCES.

GUILFORD, J. P. (1956). *Fundamental Statistics in Psychology and Education.* Third Edition. New York : McGraw-Hill.

JERSILD, A. T., and TASCH, R. J. (1949). *Children's Interests and What They Suggest for Education.* New York : Horace Mann—Lincoln Institute of School Experimentation, Teachers' College, Columbia University.

PEEL, E. A. (1948). Assessment of interest in practical topics. *Brit. J. Educ. Psychol.*, 8, 41-47.

PENNINGTON, A (1950). A study of leisure interests in schoolboys. *Univ. of Durham, Inst. of Education, Research Review*, No. 1, 31-36.

SMALLEY, L. M. (1958). A practicable diary technique for time sampling the everyday life of children. *Univ. of Birmingham, Inst. of Education Educational Review*, 10, No. 3, 233-247.

The Interests of Secondary School Children in Different Aspects of Geography.

By C. D. TIERNEY

Abstract of thesis presented for the degree of M.Ed. in the Faculty of Education, the University of Manchester, September, 1960.)

THE PROBLEM.

The purpose of the investigation was to discover whether children in secondary modern schools varied significantly in their interest in different aspects of geography.

METHOD OF MEASUREMENT.

Strength of interest in different aspects of geography was measured by a subjective-type questionnaire of forty-five times. These items were short statements taken from the fields of physical, economic, and cartographic geography. They were grouped into fifteen sets of three items, each set representing the three geographical categories. Pupils completed the questionnaire by rating each set of items in rank order of interest. Three scores were obtained for each pupil by totalling all his ratings for each category.

The questionnaire was tried out in six different forms before the above version was selected. These experimental versions used different numbers of categories, different geographical content, and different scoring methods. Item analysis was used to eliminate items which were ambiguous, or which produced responses markedly inconsistent with others in the same category. The rank order method of scoring was adopted because it was found to be slightly more reliable than scoring by either paired comparison of items or a five-point rating of each item. Various estimates of the overall reliability of the questionnaire ranged between ·82 and ·92.

The questionnaire was assumed to have content validity. A check upon validity was obtained by requiring pupils to identify the category of subject-matter to which each item belonged. A second test of validity was made by correlating scores from the questionnaire with an external criterion in the form of choice of essay title. This resulted in a validity of co-efficient of ·50.

RESULTS OF EXPERIMENTS.

The questionnaire was employed in two experiments. In the first it was applied to all pupils of one secondary modern school. A preliminary analysis of variance between total scores for each category indicated that interest varied significantly between the three categories, and that the children tested were favourable to physical and economic geography but relatively unfavourable to cartographic geography. Scores for each category were grouped into sixteen cells, according to four age groups, boys and girls, and an upper and lower ability level. The means from each cell were subjected to a three-factor analysis of variance.

The results were as follows :

TABLE 1

MEAN SCORES FROM RANK-ORDER RATINGS.

	Physical					Economic					Cartographic				
Source	s.s.	d.f.	M Sq.	F.	P.	s.s.	d.f.	M Sq.	F.	P.	s.s.	d.f.	M Sq.	F.	P.
Age	243	3	81		NS	146	3	48·7		NS	637	3	212·3		NS
Ability	885	1	885	5·9	5%	104	1	104		NS	380	1	380		NS
Sex	716	1	716	4·8	5%	827	1	827		NS	4	1	4		NS
Ability × Age	765	3	255		NS	1360	3	453·3	5·6	5%	2787	3	929		NS
Sex × Age	99	3	33		NS	356	3	118·7		NS	125	3	41·7		NS
Ability × Sex	0	1	0		NS	1702	1	1702	21·1	1%	1641	1	1641	3·3	NS
Residual	837	3	279			129	3	43			1608	3	536		
Total	3545	15				4624	15				7182	15			

The second experiment involved applying the questionnaire to a representative sample of pupils from all secondary modern schools of a large county borough. The nine secondary modern schools used were selected to reflect the proportionate numbers of children in schools of different size and type (boys, girls or mixed)

A preliminary analysis of variance of total scores for each category was made and produced a similar result to the first experiment. An analysis of variance of scores from each category identical in type with that of the first experiment was carried out with the addition of two subsidiary analyses. The subsidiary analyses were three-factor analyses of variance of mean cell scores in which only first choices were counted, and of total cell scores obtained from first choices only. Similar results were obtained from the three sets of analyses which are set out below :

TABLE 2

MEAN SCORES FROM RANK-ORDER RATINGS.

	Physical					Economic					Cartographic				
Source	s.s.	d.f.	M Sq.	F.	P.	s.s.	d.f.	M Sq.	F.	P.	s.s.	d.f.	M Sq.	F.	P.
Age	865	3	288·3	2·2	NS	131	3	43 7		NS	877	3	292·3	5·1	5%
Ability	163	1	163		NS	0	1	0		NS	182	1	182		NS
Sex	315	1	315		NS	1463	1	1463		NS	420	1	420	7·3	5%
Ability × Age	138	3	46		NS	119	3	39·7		NS	122	3	40·7		NS
Sex × Age	728	3	242·7	2·9	NS	639	3	213	7·5	5%	192	3	64		NS
Ability × Sex	81	1	81		NS	27	1	27		NS	9	1	9		NS
Residual	363	3	121			52	3	17·3			249	3	83		
Total	2653	15				2431	15				2051	15			

TABLE 3

FIRST CHOICE MEAN SCORES. FIRST CHOICE TOTAL SCORES.

	Physical		Economic		Cartographic		Physical		Economic		Cartographic	
Source	F	P	F	P	F	P	F	P	F	P	F	P
Age		NS		NS	6·4	1%		NS		NS	2·4	NS
Ability		NS		NS		NS		NS		NS		NS
Sex	4·6	NS		NS	4·5	NS		NS	3·9	NS		NS
Ability × Age		NS		NS		NS		NS		NS		NS
Sex × Age		NS	13·6	1%		NS	·26	NS	4·3	1%		NS
Ability × Sex		NS	8·2	5%		NS		NS		NS		NS
Ability × Sex × Age		—	—	—	—	—		NS		NS		NS

CONCLUSIONS.

(i) The questionnaire discriminated between groups of secondary modern school pupils in their interests in geography.

(ii) Physical and economic geography were found to be relatively popular and cartographic geography relatively unpopular.

(iii) Interest in economic geography varied with age and sex. It tended to be more popular with girls and seemed to lose the boys' interest as they grew older.

(iv) Cartographic geography proved generally unpopular but significantly more so with girls. Interest also altered significantly with age, being lowest in the second year and highest in the third.

An Analysis of the Interests of Secondary School Pupils: A Factorial Study Leading to the Construction of a Diagnostic Test

BY P. H. SANDALL,

(Abstract of a thesis submitted for the degree of Ph.D., University of London, 1960)

PURPOSE OF RESEARCH.

The purpose prompting the inquiry was the need for a test of the general interest of boys and girls of secondray school age, constructed on a scientific basis, and in a form suitable for use in school counselling and careers advisory work.

THE RESEARCH PROBLEM.

For such a test to have structural validity it was necessary that it should have a factorial construction. The theorectical part of the inquiry, therefore, was an attempt to arrive at an acceptable solution of the nature of the factors underlying the range of specific interest activities of boys and girls. The hypothesis to be tested was that these interests can be reduced to a small number of relatively independent categories at a more general level.

As a preliminary to this work, an examination was made of the psychological concept of interests, and the classifications of interests used by various research workers in the U.S.A. and in this country were reviewed.

THE SAMPLE.

The sample consisted of 429 pupils, aged from 11 to 16 years of age, attending a bilateral grammar and technical school in a semi-rural part of Sussex. Of these, 175 girls and 172 boys were in the grammar school, and 82 boys in the technical school, the range of mental ability being from 105 to 135 (Otis). About half of the pupils lived in a town of about 12,000 inhabitants, and the others in the surrounding countryside and villages; thus, both urban and rural interests were represented in the sample. There was a wide diversity in the range of parental occupations and cultural background of the pupils. As the inquiry was concerned with the relationships between kinds of interests, it was assumed that such relationships as appeared in the interests of the sample would be found also in those of the juvenile population at large. Support for this assumption was subsequently forthcoming from the testing of 2,400 pupils with the derived test, and from the evidence of other inquiries.

EXPERIMENTAL PROCEDURE.

As a preliminary step some 500 specific interest activities were examined, and the apparent relationships observed among these suggested the formulation of twenty-eight categories, as follows: Artsitic, Athletic, Children, Civic, Clerical, Collecting, Crafts, Dancing, Dramatic, Earning, Entertainment, Exploring, Handy, Health and Fitness, Humanitarian, Fiction, Indoor Games, Literary, Mathematical, Mechanical, Musical, Personal Appearance, Pets, Religious, Rural, Scientific, Sociable, Team Games.

The measuring instrument yielded individual scores for these categories, by use of the technique of forced-choice in paired comparisons. As each item consisted of the presentation of a choice between two activities representing two of the twenty-eight categories, the test consisted of 378 such items in all. The internal reliability of the test, estimated by ' split-half ' correlation, gave average product-moment co-efficients of ·72 (boys) and ·70 (girls). It may be noted that the corresponding figure obtained by Guilford for the tests used in his large-scale inquiry into adult interests in 1954 was ·72. An estimate of validity through self-rating with a sample of fifty pupils gave a coefficient of ·73.

TECHNIQUES OF ANALYSIS.

Correlation coefficients for the scores of boys and girls on the twenty-eight variables were calculated by Vernon's Double Tetrachoric Method separately for each sex, and these provided the matrices for factorial analysis by Thurstone's Centroid Method. Tests for the significance of the factors extracted gave estimates of eight for boys and seven for girls (Guilford and Lacey) or seven and five, respectively (Burt and Banks). As the proportion of variance taken out by these factors formed only about 50 per cent., and as the remaining factors tended to fall away in importance rather gradually, it was decided to proceed with rotation of nine factors in each analysis. The following solutions were selected as the most satisfactory ones. It will be noted that the factors are bi-polar.

FACTORS AND KEY TESTS—BOYS.

I.—Handy; Crafts; Mechanical; Rural; Pets; *versus* Dancing; Drama; Sociable.

II.—Indoor Games; Fiction; Team Games; *versus* Religious; Humanitarian.

III.—Team Games; Athletic; Health; *versus* Literary.

IV.—Rural; Pets; *versus* Mathematical; Clerical; Civic; Literary.

V.—Artistic; Musical; *versus* Sociable.

VI.—Civic; Exploring; *versus* Appearance; Children.

VII.—Mechanical; Scientific; *versus* Children.

FACTORS AND KEY TESTS—GIRLS.

I.—Scientific; Civic; Religious; Musical; *versus* Entertaining; Fiction; Indoor Games; Earning; Dancing.

II.—Rural; Pets; Crafts; Handy; Collecting; Earning; *versus* Dramatic; Dancing; Literary; Appearance; Entertaining.

III.—Humanitarian; Children; Religious; *versus* Team Games; Athletic; Sociable; Indoor Games

IV.—Clerical; Mathematical; Health; *versus* Artistic; Pets; (Musical; Literary).

V.—Dancing; Children; Sociable; *versus* Collecting; Literary.

VI.—Civic; *versus* Health; Mathematical.

Subsequently, the averaged coefficients of boys and girls were subjected to factorial analysis in U.S.A. by Dr. C. Wrigley, using his recently devised Quartimax method (Neuhaus and Wrigley), which, it is claimed, forms an approach to orthogonal simple structure that removes the subjective element in rotation and satisfies the criteria required by Thurstone. The calculations were made by an electronic computer to which this technique is particularly well suited. Five independent analyses were made, for 4, 5, 6, 7 and 8 factors, respectively, these being rotated to maximise the zero loadings in each case.

Interpretation of the data supplied by Dr. Wrigley showed that the results of these analyses tended to confirm those yielded by the two Centroid analyses, but were rather closer to that of the girls.

SUMMARY OF FINDINGS

1.—The results of the several analyses lend support to the hypothesis that juvenile interests can be reduced to a comparatively small number of categories at a general level The recurrence of certain tests in more than one factor showed overlap and failure to obtain complete independence in the factors. Further refinement of the measuring instrument might secure a closer approach to simple orthogonal structure, but an inherent difficulty is that sometimes more than one aspect of interests may be present in a particular activity.

2.—The factors obtained for the two sexes did not wholly coincide, although certain factors appeared to be common to both. It is not possible to assert, on the evidence of this inquiry alone, that the observed differences represent basic divergences.

3.—The results of the Centroid and Quartimax analyses suggest the presence of some of the more fundamental traits of personality, such as ' Sociable *versus* Unsociable, ' Intelligent ' *versus* ' Unintelligent,' and ' Extrovert ' *versus* Introvert.'

4 —From the results of the Centroid analyses, eight ' fields of interest ' were inferred. The contributory clusters of tests are given in brackets :

(1) Rural—Practical (Rural, Pets, Crafts, Handy).
(2) Sociable—Display (Drama, Dancing, Sociable, Appearance).
(3) Humanitarian (Humane, Religious, Children).
(4) Entertainment (Indoor Games, Entertaining, Fiction).
(5) Athletic (Team Games, Athletic, Health and Fitness).
(6) Literate (Literary, Clerical, Civic).
(7) Aesthetic (Artistic, Musical).
(8) Scientific—Mechanical (Scientific, Mechanical, Mathematical, Handy)

APPLICATION OF THE FINDINGS.

The immediate application of the findings was the construction of a diagnostic interest test, based on the eight categories given above. The twenty-eight categories and 378 items of the original test were reduced to eight categories and 120 items. The latter were selected by item-analysis from the first test as those discriminating the most within the eight categoris of the final form. Using a simple weighting, these items were combined to allow scores of from 0 to 20 in each category.

The test was standardised by administration to 1,400 boys and 1,000 girls, pupils in eleven secondary schools, grammar, modern and technical, in different parts of the country, chosen to provide a wide range of environment. To avoid undue weighting by the secondary modern pupils, the numbers from the three types of schools were nearly equal.

Product-moment coefficients for internal reliability by the split-half method were obtained from the scores of random samples of 100 boys and 100 girls, giving means of ·78 and ·74. An estimate of stability was found by retesting at intervals of three to four weeks, the mean coefficients for this being ·90. Further evidence of stability was provided by retesting in two categories at an interval of four months, during which some shift of interest may be expected The coefficients for these were ·78 and ·84. Correlation of the scores of two samples from different age-groups with their own ranking order gave coefficients of ·73 and ·70. Regarded as an interim estimate of validity this is not unsatisfactory for a personality test, but the practical validity of the test must be determined by experience in actual situations where its use is appropriate.

The scores of the whole sample were classified according to sex, age-group and type of school, and from these the tentative norms were established. Individual interest profiles could be compared with profiles representing the norms, to show at a glance strong and weak interests An as alternative, a five-fold grading, based on division at the 10th, 30th, 70th and 90th percentiles, allowed individual scores to be classified as ' Very High,' ' High,' ' Average,' ' Low,' or ' Very Low.' It was felt that the nature of the test, its breadth of categories and comparatively coarse scaling, did not warrant finer discrimination, and that these were adequate for most purposes.

SOME CHARACTERISTICS OF THE TEST.

1.—The scores conform to the normal curve of distribution.

2.—The test shows the pattern of an individual's interests under the eight broad categories, but, by reason of the technique used in its construction, cannot give any indication of differences in the total capacity for interest possessed by individuals.

3.—The stability of the test appears to be high, and it may be expected to throw some light on the permanence of interests.

4.—It reveals the expected differences between the general interest patterns of boys and girls, and shows that these are characteristic of all the age-groups and schools sampled.

5.—Where significant changes occur in the interests of different age-groups, these follow a smooth curve of increase or decrease.

6.—The interests of boys and girls from urban and rural environments are not clearly differentiated in the sample population. Whether this is due to the breadth of the categories or to an actual lack of such differences remains conjectural.

7.—There are significant differences between the norms of the three types of schools in certain categories.

8 —The presence of aversion (or marked indifference), as well as strong interest, is revealed, and it is found that these occur regularly in certain profiles, which is of particular diagnostic value in school and vocational guidance.

BOOK REVIEWS

Allport, G. W. (1961). *Pattern and Growth in Personality.* New York: Holt, Rinehart and Winston, pp. xiv+593.

It is twenty-four years since Gordon Allport published his *Personality: A Psychological Interpretation*—a book which won as wide respect in this country as in America, or wider in so far as he skilfully combined the American empirical approach to the study of people with the more theoretical European approaches. Since that time he has been continuously active in the field, and has in particular elaborated his notion that personality is not just a resultant of genetic and environmental forces acting on the organism in the past (as both S-R psychologists and the Freudians would have us believe), but is an active, self-shaping unity, the main key to which lies in the person's intentions, in what the wants to be (cf. *Becoming*, 1955).

The present volume replaces, but does much more than revise, the earlier one. Essentially, it synthesizes the original and subsequent work, and gives full recognition to the many developments that have taken place in behaviour theory, in social psychology (e.g., role theory), in personality and projective testing, and in work on person-perception, etc Not only is it a comprehensive and mature treatment, but is also beautifully written and clearly argued. Although it discusses child development and personality only incidentally, and character training scarcely at all, one can strongly recommend it to any who want to know what psychology can tell them about real people.

Anastasi, Anne (1961). *Psychological Testing* (Second Edit.). New York: The Macmillan Co., pp. vii+657, 52s. 6d.

This is the second edition of a book first published in 1954. That a second edition was required underlines its usefulness and popularity in America and stresses the changing pattern and content of psychological tests. The present volume has been brought up to date with detailed descriptions of new tests and revision of old tests. This is more than a handbook on test material available for use in schools, colleges and clinics. There is a comprehensive survey of the development of testing and a critical account of the underlying assumptions. There are useful introductory accounts of methods of expressing scores, of test reliability and the various methods for determining validity. Many additions have been made to the original text including a fuller account of the Wechsler scales, new topics dealing with test anxiety, the effects of coaching on test results and a summary of the present ideas on personality assessment.

This volume can be recommended as a most useful reference book. The descriptions and critical analyses of all types of tests give a comprehensive view of testing problems. There are perhaps more references to the results obtained by British research workers than are usually found in American text-books but references to our tests are mainly confined to Raven's Progressive Matrices and Wing's Standardised Tests of Musical Intelligence. W. H. King.

Harris, T. L., and Schwahn, W. E. (1961). *Selected Readings on the Learning Process.* Oxford University Press, pp. xi+428, $3.50.

Selected Readings on the Learning Proccess has, it is stated, been designed " to supplement text-books in educational psychology by providing ready access to source materials on the central problem of educational psychology—the learning process." Of the thirty-five readings five are definitely non-experimental, three of them being from educational texts. If the aim is as stated, that of supplementing text-books of educational psychology, these would have been better omitted and the space used for a longer introductory statement—they devote only two pages to this—of the view of learning that governed their choice and the place of each reading in their general scheme. But these foreign bodies make up only one-tenth of the text, and in an appendix the authors relate the readings by number with the relevant chapters in each of fifteen American text-books of educational psychology or books by American educational psychologists on learning. It would not be difficult to do the same for an English text-book on educational psychology.

The book includes readings on reasoning from the work of Maier and Duncker, Patrick's *Creative Thought in Poets* and Eindhoven and Vinacke's *Creative Processes in Painting*, and Hull, Smoke, and Hanfmann and Kasanin on Concept Formation. Three readings—by Book, Longwell and Vince—are devoted to the learning of motor skills, three—Sherif, Carlson and Ingersoll—to attitudinal .learning, and three—Lewin, Bales and Strodtbeck, and Barron and Krulee—to group learning processes. The remainder of the book includes readings on such topics as the development of scientific and mathematical concepts, measuring the results of college instruction, the motivational aspects of learning, the retention of meaningful materials, and the transfer value of Given and Individually Derived Principles.

Though not all the topics treated have been listed, it is clear that not only is the view taken of learning a wide one, but that the readings chosen are with very few exceptions, relevant to the kinds of learning that teachers are chiefly concerned with.

It is probably on the ground of relevance that the work of the learning theorists and work on non-human animals have been entirely excluded ; this, though perhaps surprising in an American book, is both logical and resfreshing. There is a corresponding absence of those implausible, uncritical and unproven accounts in terms of leaning theory of the sort of learning that occurs at home and in school, in which books based on learning theory abound.

One feature of the book will strike every non-American, that is, with the solitary exception of Piaget, the readings are all from either Americans or European immigrants to America. Still, even with this limitation, the book is to be strongly recommended for use in the way described. The relevance of the readings to teaching, the preponderance of the more developed forms of learning, and the catholicity of the selection make this a valuable addition to the growing number of collections of supplementary readings for psychology students. H. E. O. JAMES.

KLAUSMEIER, H. J. (1961). *Learning and Human Abilities: Educational Psychology*. New York : Harper, pp. x+562, 52s.-6sd.

This is a text-book for students. It is a study of classroom teaching, very thoroughly done. All aspects of the school situation are analysed psychologically, relevant experimental work is fully presented and discussed, and the student is given advice on how to deal with situations which may arise.

The book contains a great deal of immensely useful material, but it is hard to estimate how far British students are likely to find it usable. The bulk of the book and the quantity of material may discourage the less academic students, who may also find some of the theoretical analysis too much for them. Graduates may find the book unnecessarily slow and specific. There is a tendency to tell the student in detail, and often more than once, what to do in particular situations ; rather than to enlarge his general understanding so that he can see what to do for himself. Sometimes one feels the reader is being ' taught ' as if he were a child : " To clarify the relationship between capacity and ability, imagine two quart milk bottles side by side. One bottle is half full, the other entirely. The present ability of the two bottles to provide milk is unequal, although both have the same capacity."

The style of the book may be an obstacle for some readers. It combines superficial briskness with an inexact prolixity which makes it tiring to read. Sometimes it is actually obscure : " Answers which the child wants to solve an imporrant problem are deferred because the problem which an adult wants solved takes precedence." The book is probably not exactly suited to any category of British students, but those who do find it readable may also find it very valuable.

S. C. RICHARDSON.

Lindgren, H. C., and Byrne, D. (1961). *Psychology: An Introduction to the Study of Human Behaviour.* London: Wiley, xi+429, 52s.

The authors state that the central purpose of writing the book is that of aiding psychology students to develop a better and more complete understanding of human behaviour. It certainly covers a wide field. After an opening chapter in which Psychology as a Science and Profession is discussed, there are two chapters dealing with many different facets of Human Development from Infancy to Adulthood. This is followed by a chapter on Physiological Processes, one on the Learning Process—in which the problem is looked at from a number of viewpoints—while Chapters 6 and 7 deal with Perceptual Processes, and Motivation and Emotional Behaviour, respectively. Then follow discussions on the use of Statistics in Psychology, Intelligence and Creativity, and Personality. The last seven chapters deal with Personal Adjustment and Mental Hygiene, the Social Matrix of Individual Behaviour, Communication, Group, Processes, Selecting and Learning Occupational Roles, Psychology in Business and Industry, and Psychology and World Affairs.

Apart from covering a wide field with up-to-date references, the work is written clearly and simply at an elementary level. There are good diagrams; also questions, and suggestions for further reading at the end of each chapter, while at the end of the book there is a glossary of technical terms.

For the British or European student, however, the book has some disadvantages. First, the text is written wholly in terms of the American social setting. Second, it almost ignores European psychology. The names of Burt, Broadbent, Eysenck, McDougall, Piaget, Valentine, the Vernons, Zangwill and a host of other well-known workers do not appear at all. There is one reference each to Binet, Bowlby and Tinbergen, and two references to F. C. Bartlett. Third, photographs are found on over eighty pages; some of these are of great value, others tell the reader nothing at all (e.g., the photographs on pages 105, 175, 293). Again, there are humorous line drawings on over sixty pages, most of which communicate little or nothing of value. The British and European reader would probably prefer a drastic pruning of the photographs and line drawings, thereby enabling the price and size of the book (the page measures 10½-in. by 7½-in.) to be reduced, or important contributions from this side of the Atlantic to be included.

Nevertheless, the British reader could gain much from a study of this book whether or not he is a student of psychology. K. Lovell.

Lundin, R. W. (1961). *Personality: An Experimental Approach.* New York: Macmillan, pp. xiii+450, 40s.

Dr. Lundin's book attempts to provide a basic psychology of personality wholly in terms of Pavlovian and Skinnerian conditioning Traits, motives and other 'intervening variables' are ruthlessly discarded. Almost all the chapters begin with rat experiments, and then go on to show how similar are the social and personal reactions of normal and maladjusted humans to those of rats. At the same time, Lundin draws heavily on Freud who, like Skinner, tends to account for present behaviour largely in terms of influences of early childhood; and it is, indeed, striking how many of Freud's 'mechanisms' have been duplicated among rats.

Clever as the book is, it hardly seems to prove that a purely Behavioristic psychology of personality is adequate. After all, a close friend of Dr. Lundin's, or a clinical psychologist, who employed such unscientific concepts as traits, attitudes, etc., might be able to predict the sort of wife that Dr. Lundin would marry. But it seems most unlikely that Professor Skinner, by studying Dr. Lundin's previous exposures to variously reinforced stimuli, could make an equally good prediction Nevertheless, one should not merely dismiss the author's approach with ridicule. He is surely justified in claiming that people acquire their personal dispositions in accordance with known principles of learning; and that therefore, through further studies of learning, we should, in time, be better able to build desirable characters or cure the maladjusted than we can at present through rule-of-thumb or through largely unscientific clinical procedures. P. E. Vernon.

MAXWELL, A. E. (1961). *Analysing Qualitative Data.* London: Methuen, pp. 163, 16s.

Dr. Maxwell's new book is an admirably clear and succinct summary of the main techniques involving the use of chi-square for the analysis of data classified in categories. It includes Fisher's exact test for small samples, treatment of matched samples, partitioning, goodness of fit, and other more advanced techniques. There is a useful chapter on rho, tau and W for ranks; another on the sign test and Cox's test for trend, and a final—not very enlightening—one on item analysis.

However, the book is rather less useful for educational researchers (whose data are more often continuous) than for those working in the clinical or medical fields, from which most of the illustrations are drawn. Also it is rather advanced for ordinary students. For example, the author does not even bother to warn against their most besetting fault, namely applying chi-square to non-independent events, such as numbers of delinquencies committed, or books read, by certain groups of children. Nor does he advise on the tricky question of interpreting results when large numbers of chi-squares have been calculated for the same groups.

MUSSEN, P. H. (Edit.) (1960). *Handbook of Research Methods in Child Development.* New York: John Wiley, pp. x+1061, 122s.

There are probably far too many secondary source books on psychology being published in America nowadays, including annual reviews, readings, symposia and multi-authored manuals; though admittedly it is difficult for lecturers, let alone-students and practitioners, to keep up with primary publications. However, a good case can be made for the present Handbook, even if its weight of nearly 4 lbs. belies its name. It definitely fills a need not met by Festinger and Katz, or by Carmichael's well-known Manual. Obviously, few students could buy it, but it is a 'must' for education department, and we hope training college and teacher, libraries.

It could, in fact, have been cut to half the size and been much more usable had it been written by a single author, or a small group. There is far too much repetition. Thus, time and again, the same contribution is described from slightly different viewpoints in the chapters on Interviewing (L. J. Yarrow), Motivation (D. R. Miller), Projective Techniques (W. E. Henry), Personality Appraisal (U. Bronfenbrenner and Ricciutti), Attitudes (M. R. Yarrow) and Family Life Variables (L. W. Hoffman and R. Lippitt). However, the indices are excellent, and the research worker in child psychology should be saved many blunders in future by following up what half-a-dozen or more chapters have to say about his strategies and teachniques Another weakness is the over-concentration on the young child in the home or nursery school. Only a fraction of the space is devoted to studies in the primary school, still less to the secondary, or to children in the out-of-school-or-home environment. Readers in this country may bridle, too, at the virtual neglect of any European contributions to child psychology except Piaget's (though a few sops are thrown to Bowlby, Isaacs and one or two others). But this is not due merely to American provincialism. Most of the important and sophisticated developments in methodology are fairly recent, and are attributable to American authors. Nevertheless one would have hoped that Burt's work on delinquency and Himmelweit's study of television would be worth mentioning.

For reasons of space we can do little more than list the constituent chapters and authors. A. L. Baldwin and W. Kessen give general surveys of problems in child study, and provide much sound advice on ethical considerations in procuring and motivating subjects H. F. Wright on Observational Methods, Bijou and Baer on Experimental Techniques and Yarrow on Interviewing are also thoroughly useful. There are good accounts by specialists of Physical Growth (H. V. Meredith), Chemical and Physiological Development (I. G. Macy and Kelly), Sensation (A. H. Reisen), and Perception (E. J. Gibson and Olum). C. C. Spiker is all right on conditioning and discrimination learning, but he and the Editor seem not to have noticed that children also learn at school. B. Inhelder and Matalon cover Piaget's methods of studying thinking, and a lot more besides.

Readers' interests will differ, of course, but the present reviewer found J. Berko and R. Brown on Psycholinguistics, and J. W. M. and B. B. Whiting on anthropological techniques particularly exciting. A. Anastasi on Ability Tests and the aforementioned chapters on personality tests are all good in their way, as is G. C. Thompson on Children's Groups. W. W. Lambert's highly theoretical treatment of Interpersonal Behavior seemed somewhat out of place; but the final chapter on Family Life Variables (influence of social class, parental attitudes, etc.) rendered much of the rest of the book redundant. P. E. VERNON.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1961). *Studies in Spelling.* University of London Press, pp. xi+208, 15s.

Singularly little work on spelling has been carried out by British educational psychologists, other than Schonell, and it is left to a group of Scottish teachers to produce the first full-sized book on the subject. The efforts of this Panel, set up by the Scottish Research Council, have been very well worth-while. They have already produced a new *Scottish Pupils' Spelling Book* by methods described in the volume under review, where the choice of words was based on a large sampling of junior school children's written usage.

The Convener of the Panel, H. J. L. Robbie, contributes a fascinating essay on the history of English spelling—its diverse origins and the many influences which have led to its gradual standardisation. He points out wisely that standardization is far from complete (an example occurs in this and the preceding sentence); and that some flexibility and further modifications are to be welcomed. Later, Dr. W. Boyd outlines a possible approach to simplified spelling. The most substantial section is a review by A. C. Paterson of the research literature on spelling, running to about 100 references. This points rather strongly to the desirability of teaching spelling through writing rather than as a separate discipline, together with the superiority of the 'Test-Study' to the conventional 'Study-Test' approach. Also included are accounts of three well-conducted Ed.B. researches on aspects of spelling teaching.

STEVENS, F. (1960). *The Living Tradition: The Social and Educational Assumptions of the Grammar School.* Hutchinson and Co., Ltd., pp. 304, 35s.

This book is haunted by a living tradition of quite a different kind from that referred to in the title, that of the working class world It is an account of the attempts, and refusals, of both pupils and staff to come to terms with what—as one headmaster said— is "a matter of communicating middle-class values to a 'new' population."

The method used by Miss Stevens in her book is primarly anecdotal; from this, it derives its force, charm and immediacy, although conscientious investigations were carried out and there are a number of diagrams and graphs at the back of the book. Miss Stevens collected questionnaires from nineteen schools and her book is based on the answers of 120 teachers and over 2,000 pupils. She also reports conversations with sixth formers and staff of twenty schools, 'not exactly identical,' with the questionnaire schools.

The conversations appear in one part of the book; 'The Teachers Speak,' through questionnaires, in another; 'Those who are Taught' in yet a third. This concession to systematisation is a mistake for there is little attempt to give examples of the inter-action of pupils and teachers: no school is presented as a unit. Nevertheless, many important problems are aptly illustrated. They need to be pinpointed by further investigations and interpreted in the light of psychological knowledge—as, for example, "the curious fact that whatever the rate of selection the estimate of the pupils in the school capable of starting Latin was roughly constant." MILDRED COLLINS.

SYMONDS, P. M., with JENSON, A. R. (1961). *From Adolescent to Adult.* Columbia University Press. London: Oxford University Press, pp. viii+413, 70s.

This book is a follow-up study of the girls and boys who were the subjects of the late Professor Symonds' volume *Adolescent Fantasy*, with which he introduced his Picture-Story Test. However, this new publication makes sufficient reference to the former data to stand alone, and it is of wider scope than that of projective psychology only, being also a social document of great interest.

For the historian looking at the growing integration within psychology, " much has transpired in thematic apperception testing theory over the thirteen intervening years, and there is closer rapprochement between testing theory and psychodynamic theory in general " (p. 27) Of particular interest is the addition of the Rorschach as a research tool complementary to the Picture-Story and interview material (pp. 119-175), with acknowledgment of validation difficulties. This central part of the book is typical of the author's admirable skill in combining quantitative and qualitative approaches

The value of the book for many people will lie in its contribution to the psychology of development Chapters 3, 5 and 11 for example, consist of comparison of adolescent with adult fantasy; correspondence between adolescent fantasy and subsequent experience, relationship between adolescent and adult personality. The authors consider that one of the outstanding findings is the remarkable persistence of personality trends over the thirteen-year period. One of the questions which it was hoped to answer was whether it is possible to predict mental health from the kind of data gathered in 1940. A correlation of ·54 was obtained between the estimates of adjustment made then, and again in 1953. The next chapter deals with some conclusions on the role of fantasy in personality. " Fantasy is the ' unfinished business ' of personality, that is, it deals with unexpressed wishes and desires, and unresolved conflicts. Personality is a precipitate, a working out, and a resolution of fantasy," This is an arresting thought, and although probably not intended as the whole story, it goes some way towards giving a rationale for the persistence of personality and the fairly high level of prediction. The changes over the years are as interesting, of course, as the similarities, and their relationship is seen in Appendix B, the case of Jack, who was studied in detail in *Adolescent Fantasy*. Appendix A gives a resume under six headings of the main facts about each subject.

It is rare to find such unselfconscious, readable documentation of the details of a research process. This important book gives a vivid impression of colleagues bringing complex psychological material within the scope of scientific treatment without loss of human reality and with the utmost integrity. As such, it would make good reading for research workers in any field. L. R BOWYER.

TORRANCE, E. PAUL (Edit.) (1961). *Talent and Education. The Modern School Practices Series, No.* 4. Minnesota University Press. London: Oxford University Press, pp. x+210, 36s.

The recent interest in the United States in the search for talent in schools and colleges has given rise to a number of interesting researches and stressed the consideration of problems dealing with the nature of abilities, the importance of the influence of the environment and the need for personality adjustment This volume is the combined work of fifteen contributors to the 1958 Institute of Gifted Children sponsored by the Bureau of Educational Research, University of Minnesota. A symposium of this nature covers a wide variety of interesting questions and a few, perhaps, are specific to the American background. The thoughts, ideas and research of several of these papers however, are directly relevant to our own situation at the present time. The general feeling that prompted these investigations is perhaps best summarised by F. N. Brumbaugh in her paper on " Special Grouping." " The need for manpower with above average intelligence, by industry and government, in mathematics, science, linguistics, has done more to focus attention upon the needs of the gifted than the years of effort by educators to do the same."

Perhaps we fight shy of the term ' gifted ' as applied to the top 5 per cent. or 10 per cent. of the ability range ; perhaps we think that any problems we have in this field are solved by the academic achievements of the grammar schools and the standard of entry to our universities. Everyone concerned with junior and secondary education will find interesting the arguments set forth for enrichment of the school curriculum, for special grouping, and for acceleration and for the formative role of experience and environment.

D. B. Harris approaches (as a psychologist), the problems raised in many of the contributions. He emphases the lack of agreement on whether the gifted should or should not be dealt with in separate classes all the time or part of the time, and as other contributors do, points out the vital importance of a gifted teacher and the importance early on in the child's life of the family providing rich stimulation and suitable attitudes and values. Psychometrists will be interested in one of his statements as evidence of the narrowing of the differences between British and American opinions. " . despite the emphasis on the factorial approach in recent years, and the discussion of abilities rather than ability, there is accumulating evidence on some ' g ' factor, at least substantial positive intercorrelations in clusters of abilities . . we are not faced with a welter of specifics, each of which may require a special approach "

The contribution of John E Anderson on the problems relevant to the development of abilities expresses much of our own thought at the present time. His reference to true, measured and effective ability should be compared with the interpretation over here of Hebb's approach to intelligence Anderson's insistence on the necessity to provide young children early with stimulating environment agrees with the published findings of Piaget's work

The research report by Paul Torrance on under-achievers and their self-concepts might well be a subject for further research here He finds that constant movement of the home in childhood produces a peculiar pattern of isolation and those of high ability protect themselves in new situations by under-evaluation. No review of this symposium can ignore the very informative and objective report of R. H. Beck on his recent visit to the U.S.S.R. He outlines the Russian philosophy and its bearing on educational ideas ; the role of the teacher in the school and the standards of achievement. He describes the Russian concentration on the type of environment provided in and out of school and the reasons for the complete denial that there are any " inherited limits on the ability to learn " or that there is any recognition, the schools, of individual differences. W. H. KING.

VINCENT, E. L., and MARTIN, P. C. (1961). *Human Psychological Development.* New York : Ronald Press, pp. v+522, $6.50.

The authors tell us that this book grew out of a number of years of teaching a general education course to freshmen and sophomores in a liberal arts college. Their aim is to give students a better understanding of human psychological development both in themselves and other people, and they presuppose no previous knowledge of psychology.

After an introductory survey of the main primary and secondary influences on development and behaviour, there follow a series of chapters dealing with the successive periods of life from conception to old age. The chapters on the adolescent and early adult periods are the fullest since these are likely to be of greatest interest to the students, and those dealing with later adult life are the briefest and attempt mainly to show that this is just as much a stage of development as earlier ones and brings its own problems which individuals must solve for themselves. There is also a chapter on Courtship, Marriage and Family Life, subjects which are usually of considerable interest to students. One unfortunate result of the simplicity of the treatment is that categorical statements are occasionally made which, while not strictly untrue, need to be modified if they are to accord with all the evidence available. This defect should be realised by teachers who use the book with students.

The emphasis throughout is on the desirability of the individual's tackling the problems he meets in the course of this development, so that he may become a mature and well-adjusted adult, physically, intellectually, socially and emotionally. The treatment is admirably suited to young students. It is objective in that it directs attention to observation of children and other young people, but at the same time gives information which will illumine the students' own problems and set them on the way to finding good solutions for themselves. It is recognised clearly that everyone has to face difficulties on the road to maturity and that these are not usually unique, and pointers are given as to the best ways of dealing with them. Although it is not a work of profound scholarship, this is a sensible, kindly book, and students should be encouraged to read it for themselves. The price makes it impossible to suggest that they should buy it, but it should find a place in college libraries.

It may be advisable to point out that the book is American and that the statistics quoted refer to the United States population and should not be taken as applicable unmodified to social conditions in the United Kingdom. Also one might question whether the pictures have any real function. Their omission might reduce the cost without reducing the usefulness of the book, and this would be an advantage. These, however, are minor criticisms, and should not obscure the real value of the work. K. M. EVANS.

WERTHEIMER, MAX (1961). *Productive Thinking.* Edited by Michael Wertheimer. London: Tavistock Publications, pp. xvi+302, 28s.

Thinking is the hardest of all mental processes to represent. Models for it are not readily found and the Gestalt theory still offers one of the most workable sets of general concepts for the descriptive analysis of thought processes. Gestalt principles seem especially serviceable for making the distinction, which is the theme of Wertheimer's book, between 'productive' thinking and inefficient, 'ugly' thinking. He describes the productive thought process as a re-centering and re-structuring of the mental field, and its final product a perfect pattern of necessary, interdependent parts. 'Ugly' thinking is insufficiently re-centred, imperfectly re-structured, contains fortuitous connections and arbitrary components.

The distinction between the two kinds of thinking is perhaps not so sharp in ordinary life as Wertheimer was able to make it in some of his experiments; for instance, with the children's geometry class. When one reads his analysis of the discoveries of Galileo and Einstein, one may wonder how much 'ugly,' hit-or-miss thinking contributed to the final elegant solutions; and, indeed, whether important discoveries are ever made without a good deal of preliminary non-productive effort. But, perhaps, this is unimportant since Wertheimer does not value productive thinking simply because it gets results He feels, rather, that the productive processes are valuable in themselves, aesthetically and psychologically, and it is for this reason that he urges the need to cultivate them in children and to plan education so that 'ugly' thinking is reduced to a minimum.

The debt of modern education to the Gestalt theory is perhaps not always fully acknowledged. Wertheimer says of his own educational suggestions: "But I may say at once: there are good teachers who do similar things out of genuine feeling for what is required educationally. Most of what I shall propose is by no means surprisingly 'new.' But it is certainly in many ways different from the actual procedures in many schools" That these schools are becoming fewer, and that Wertheimer's proposals are becoming less new, is partly due to the influence of Gestalt writers on modern educational thought.

Three new chapters and six appendices found among the author's papers, have been added to the original book. The new material includes further discussion of the learning and teaching of arithmetic; the analysis of a bridge problem; and a funny story about a school inspector. S. C. RICHARDSON.

A new guide to published tests and supplement to The Mental Measurements Yearbook *series . . .*

TESTS IN PRINT

A Comprehensive Bibliography of Tests For Use in Education, Psychology and Industry

Edited by Oscar Krisen Buros

Professor of Education and Director, The Institute of Mental Measurements, Rutgers, The State University.

This new reference work will save persons responsible for selecting and using standard tests hours of labour in locating information about tests of all kinds—achievement tests, intelligence, scholastic aptitude and mental ability tests, projective and non-projective personality tests, vocational interest, aptitude, and selection tests, and numerous others. Provides information on tests originating in the United States as well as other English-speaking countries ; 149 tests originating in the United Kingdom are listed, as are 77 originating in Canada and Australia.

Through complete cross references *Tests in Print* also serves as the first available master index to the wealth of additional information and critical reviews of specific tests to be found in the volumes of *The Mental Measurements Yearbook series.**

There are also brief listings of more than 800 out of print tests (included for the convenient reference of yearbook users and students of testing) and other materials of interest to test users. The convenient indexes include directories of test publishers and overseas test distributors.

Published November, 1961. Pp. xxix, 479. Bound in cloth. $7.00 plus postage. (publisher pays postage on prepaid orders).

Order directly from :

THE GRYPHON PRESS

220, Montgomery Street, Highland Park, New Jersey, U.S.A.

* Also available from GRYPHON : *The Fifth Yearbook*, 1959, pp. 1292, $22.50 ; *The Fourth Yearbook*, 1953, pp. 1163, $18.00 ; *The Third Yearbook*, 1949, pp. 1047, $12.50. The set $50.00.

CONTENTS

PAGE

B. W. Sumotirto: Social Attitudes Among High School Students in Indonesia 3

O. C. Sampson: Reading Skill at Eight Years in Relation to Speech and other Factors 12

G. E. R. Burroughs and J. N. Morris: Factors Involved in Learning a Simple Musical Theme . 18

I. R. J. Petrie: Residential Treatment of Maladjusted Children: A Study of Some Factors Related to Progress in Adjustment . 29

M. E. Hebron: A Factorial Study of Learning a New Number System and its Relation to Attainment, Intelligence and Temperament 38

R. L. Curry: The Effect of Socio-Economic Status on the Scholastic Achievement of Sixth Grade Children 46

D. R. Price-Williams: Abstract and Concrete Modes of Classification in a Primitive Society 50

L. E. Keating: A Pilot Experiment in Remedial Reading at The Hospital School, Lingfield, 1957-60 62

K. Lovell, E. Johnson and O. Platts: A Summary of a Study of the Reading Ages of Children who had been given Remedial Teaching 66

Critical Note:

R. R. Dale: Comments on an Article by M. L. Kellmer Pringle . . 72

M. L. Kellmer Pringle: A Reply 73

Research Note:

C. D. Hardie: Note on the Problem of Inference in Educational Research 75

Summaries of Researches Reported in Degree Theses 78

Book Reviews 93

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education once a term, in February, June and November. The subscription is 30s. a year post free, payable in advance, or 15s. each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36, Essex Street, Strand, London, W.C.2." Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

The British Psychological Society also issues quarterly the *British Journal of Psychology* and the *British Journal of Medical Psychology*. Subscriptions, 60s. per annum for either journal, should be sent to the Cambridge University Press, Bentley House, Euston Road, London, N.W.1. Papers for publication in the former should be sent to Dr. Boris Semeonoff, The University, Edinburgh, 8; those for publication in the *British Journal of Medical Psychology* should be sent to Dr. Joseph Sandler, 96, Portland Place, London, W.1.

The British Journal of Statistical Psychology is issued twice a year, in May and November; Subscription 30s. 6d. per volum[illegible]st free) should be sent to Taylor and Francis, Ltd., Red Lion Court, Fleet Street, London, E.C.4; or [illegible]ny bookseller. Members of the British Psychological Society should send subscriptions (30s.) to the Secretary, British Psychological Society, Tavistock House South, Tavistock Square, London, W.C.1. Papers for publication should be sent to John Whitfield, Dept. of Psychology, University College, Gower Street, London, W.C.1.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Methuen and Co., 36, Essex Street, Strand, London, W.C.2.

Published by

METHUEN AND CO., LTD., 36, Essex Street, Strand, London, W.C.2.,

for the British Journal of Educational Psychology, Ltd.

(Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.)

on behalf of **The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.**

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS., LTD., 42, HILL STREET, BIRMINGHAM, 5.

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

JUNE 1962

VOL. XXXII, PART 2

Price Fifteen Shillings

THE

BRITISH JOURNAL

OF

EDUCATIONAL PSYCHOLOGY

Edited by

L. B. BIRCH

Associate Editors:

P. E. VERNON and DORIS M. LEE

Vol. XXXII, Part II

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is 30s. a year post free payable in advance, or 15s. each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36, Essex Street, Strand, London, W.C.2." Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Methuen and Co., 36, Essex Street, Strand, London, W.C.2.

Published by
METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of
The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

CALCUTTA UNIVERSITY
CENTRAL LIBRARY.

Dr. ARNOLD GESELL, 1880-1961

By LOUISE BATES AMES, Ph.D.
(*Gesell Institute of Child Development.*)

Dr. Arnold Gesell began his work at Yale in 1911. His purpose, strictly adhered to for the next fifty years, was to determine and describe the patterned changes of behavior which took place in the human organism as it developed from infancy through adolescence.

Dr. Gesell's basic thesis was that developing human behavior has as much structure and shape as does the developing physical organism, that this behavior develops in a patterned, highly predictable manner, and that it is measurable. According to him, " Behavior has shape. Mind manifests itself."

He emphasized especially his belief that " environmental factors support, inflect and modify, but they do not generate the progressions of development." At no time did he relinquish his basic belief that the scientist must know much more than he does about the common stages through which the behavior of all developing organisms develops, and also about the basic individuality of the organism developing, *before* he can properly evaluate the effect of the environment on that organism.

Gesell's work met with much opposition both from Behaviorists, and from colleagues of a Freudian persuasion who, in their own separate ways believed that the organism can be studied only as it is affected by its environment. However, the success of his efforts is suggested by a review by Helen Puner in *Harper's Magazine* for March, 1956.

According to this reviewer, " Now in the late fifties, constitutional factors and the nature of the baby are being mentioned and written about once again . . . It is beginning to be recognized that Dr. Gesell was correct in his belief that ' personality cannot be understood as a direct reflection of its shaping force . . . ' In fact, it may easily be that Dr. Gesell's aim—if not his point of view—has not been so much behind the times as ahead. For to-day, it is he and he alone who has built up the only systematic body of observations,on the development of human behavior and personality."

Dr. Gesell's unusual productivity during a long career (he worked actively as research consultant at the Gesell Institute which bore his name and was founded in his honor in 1950 after his retirement from Yale, until he was 78 years of age) is indicated by the length of his bibliography. This bibliography numbers well over 400 items, including books, articles, monographs and films.

Though his scientific writings covered many fields including retardation, cerebral palsy, cretinism and mongolism, visual deviation, twinning as well as behavior of normal infants and children up through 16 years of age, the largest portion of his published works dealt with infancy and the pre-school years. These books include, notably, *The Mental Growth of the Pre-school Child* (1925), *Infancy and Human Growth* (1928), *Infant Behavior* (1934), *Psychology of Early Growth* (1938), *Atlas of Infant Behavior* (1934), *Developmental Diagnosis* (1941) and *Infant and Child in the Culture of Today* (1943).

His detailed and systematic descriptions of the development of all kinds of early behavior in the human infant and young child, and the behavior tests based on his findings, were used by psychologists and pediatricians alike. In fact, it was largely due to his insistence that " development as well as disease lies clearly in the province of clinical pediatrics " that in the early 1940's the

American Academy of Pediatrics changed its admittance examination to include a section on the development of infant behavior.

In his last ten years at Yale, Gesell and his staff trained pediatricians not only from all over the United States, but from all over the world. The Yale Clinic of Child Development was known not only for its research, publications and clinical service, but also for these students who now staff clinics in many parts of the world.

Today, many pediatric clinics routinely give the Gesell Developmental tests to infants and pre-schoolers in an effort to determine not so much their intelligence as such as their D.Q. (developmental level or developmental quotient). Dr. Gesell considered that the developmental examination was, in fact, a neurological examination and that especially in cases of defect and deviation it could give a valid indication of the level of behavior development.

Pursuing his study of patterned behavior changes in developing behavior, Gesell and his staff continued their investigations beyond the pre-school years up through sixteen years of life. These findings were reported in *The Child from Five to Ten* (1946) and *Youth: The Years from Ten to Sixteen* (1946) which, along with *Infant and Child in the Culture of Today*, constituted his best-known and widely translated triology.

Much of Gesell's infant work was carried out by means of cinemanalysis, a method which he did a great deal to further. In 1925, he began his collection known as the Yale Films of Child Development. Filming was done in the now well-known Dome, and miles of systematically recorded film, both 16 and 35mm. were gradually made available for the analysis of behavior patterns and for editing into teaching and demonstration films. Dr. Gesell was also one of the first to use the one-way screen method of observation.

The usefulness and effectiveness of the developmental point of view is indicated by the fact that it has turned out to be uniquely useful not only in the fields of child psychology, pediatrics and education, but also in fields which were scarcely thought of when Gesell began his work—such widely divergent fields as those of infant and child vision, and of the projective techniques—two areas in which he did considerable effective work toward the end of his long career.

Dr. Gesell's primary interest was the scientific study of human behavior—his aim, to establish the fact that human behavior does develop in a patterned orderly manner through stages which are alike in quality and sequence (even though not necessarily in exact timing) from child to child. However, after some twenty years of systematic and controlled study of early behavior patterns observable in normatively controlled test situations, his observations broadens out to include less controlled behavior as observable in home and school situations.

This led to an extremely wide use of his findings in both homes and schools as well as in more controlled laboratory and clinical situations. At the time of his death, an editorial in the *New Haven Register* summarized this influence as follows:

" In countless American homes the name of Arnold Gesell was better known than that of the President of the United States... His work over many years in the fields of pediatrics and child development paved the way to health and success for thousands, mayhap millions. Dr. Gesell was a pioneer, one who traced uncharted paths to charted conclusions."

THE DEVELOPMENT OF FORMAL THINKING IN VERBAL COMPREHENSION

BY DUNCAN CASE AND J. M. COLLINSON
(University of Birmingham)

SUMMARY. This paper describes the results obtained by the investigators in their efforts to establish whether the inferences that secondary modern school children and pupils in further education draw in history, geography and literature texts correspond to the formal operational stage of reasoning as postulated by Piaget in his theory of mental development. Further, it was sought to provide answers to the following questions :

(i) Do the results obtained correlate more closely with mental or chronological age ?

(ii) Is there any significant sex difference shown in the results ?

(iii) Do the results show any significant difference by age ?

I.—INTRODUCTION.

MUCH work has been done, and is still being done, in the field of the pupil's thinking by staff and research students in the Education Department at Birmingham University. The research described here was inspired by Professor E. A. Peel,[7] Head of the Department of Education at Birmingham University, and carried out along lines suggested by Lodwick[6] in his attempt to test Piaget's[5] hypothesis of the three-stage development in children's logical judgments.

TABLE 1

Age	Stage	Thinking
0—2 2—7	Sensori-Motor. Intuitive, Pre-operational (irreversible).	Fragmentary partial, inconsistent judgments. Tied to one aspect of the problem.
7—11+	Concrete operations (reversible)	Tied to the concrete, tangible, manipulable present. Deals with reality.
12+	Formal operations.	Propositional, hypothetico—deductive dealing in the realm of possibility.

II.—SELECTION OF GROUP FOR INVESTIGATION.

The intention at the outset was to select a group of eighty children, ten from each age group from 11+ to 18+, divided equally between the sexes. The 11+ to 14+ group were to be secondary modern school children and the 15+ to 18+ group from a college of further education. It was also hoped that each age level should represent as wide a scatter of intelligence as possible. The intention had to be modified since the further education group covered 15+ to 17+ only with one subject 14 : 10 and another of 18 : 1. It was not possible to obtain an equal division of sexes, neither was it possible to obtain a sufficient number of subjects of low intelligence to balance the high intelligence group. A primary school group of ten children, 7 to 10 years, was included as an aid to the validation of the texts used in the investigation. A wide scatter of intelligence was obtained within this primary school group. The ninety selected subjects were given the Raven's[8] Progressive Matrices Test (1938), Sets A, B, C, D, and E, in order to obtain their Mental Ages.

III.—SELECTION OF TEXT MATERIAL.

Five texts, two history, one geography and two literature were selected for use, with three questions set for each text. The test was administered to a pilot group of ten children, within the age ranges decided. Some modification of questions and interviewing technique was indicated as a result.

The test was put to the selected group as follows: each child was interviewed separately and after an initial settling-down period of small talk and dummy questions, and answers, the child was asked to read the selected text silently. Next, the text was read aloud by the interviewer, following which the child was asked each question in turn and gave oral answers which were sometimes expanded by supplementary questions. One of the texts, with specimen answers and scores, is given in the Appendix.

IV.—TEST MARKING.

The levels of thought were quantified by dividing into three sections (as shown below) and a score for each answer was awarded which could be set against Chronological Age and Mental Age:

Stage III.	Formal Operations	3 marks.
Stage II..	Concrete Operations	2 marks.
Stage IIa.	Intuitive Pre-operational	1 mark.

Thus, the maximum score per answer was 3 marks, per text 9 marks, and per test 45 marks.

V.—RELIABILITY OF MARKING.

The answers of the whole group were marked independently by the investigators and Dr. R. M. Beard. Even the two most discrepant markers yielded a correlation coefficient of ·98. N=1,350 (90 children × 15 questions each).

Using Guttman's[4] Scalogram Technique, the scores obtained on each item were set against Total Scores, Mental Age and Chronological Age, and a Coefficient of Reproducibility was found for each.

$$\text{Coefficient of Reproducibility} = 100 \times \left(1 - \frac{\text{No. of errors}}{\text{No. of items} \times \text{No. of subjects.}}\right)\%$$

TABLE 2

TABLE OF CO-EFFICIENTS OF REPRODUCIBILITY AS PERCENTAGES FOR THE WHOLE TEST GROUP (ninety Children).

(*a*) *Tests.* (*b*) *Test as a whole.* (*c*) *Groups.*

		Text	M.A.	C.A.	Total Scores
(*a*)	History	A	55·2	55·5	67·0
		B	53·3	50·8	64·8
	Geography	C	74·8	73·3	80·4
		D	69·6	68·5	78·5
	Literature	E	77·0	78·1	77·5
(*b*)	Test as a whole	—	66·0	65·3	73·7
(*c*)	11+ to 14+ Groups	—	55·3	54·5	69·5
	14+ to 17+ Groups	—	75·1	76·2	82·51

From the coefficients it may be assumed that the group had more difficulty with the two History texts, A and B, than with the Geography text, C, or the two Literature texts, D and E. However, the results for the younger group would have us link Literature text D with History texts A and B in difficulty. We may perhaps assume that the difficulties experienced by the younger group with Text D are overcome by the Chronological and/or Mental Age of 15 years.

No coefficients for the whole group nor for the 11+ to 14+ group (but some coefficients for the 15 to 17+ group) reached the 85 per cent. standard for a perfect scale set by Guttman. This suggests that another factor or factors besides Mental Age and Chronological Age is contributing to the development of formal thought.

All the tables show close similarity between Mental Age and Chronological Age coefficients. We may conclude from this that neither is more important than the other in the development of thought structures.

TABLE 3

RANK CORRELATIONS OF TOTAL SCORES AGAINST MENTAL AGE AND CHRONOLOGICAL AGE.

All		11+—14+		15—17+	
M.A.	C.A.	M.A.	C.A.	M.A.	C.A.
·65	·66	·61	·31	·36	·36

Using Fisher's[3] Z conversion tables no significant differences were found between the correlations, though the correlation of Mental Age with total scores for the 11+—14+ group was much lower than its Chronological Age counterpart. Although all the correlations were positive they were rather low, which might again suggest the existence of another factor or factors in the development of formal thought.

VI.—ANALYSIS OF VARIANCE TO TEST SEX AND AGE.

TABLE 4

Text	Group	s.s.	d.f.	m.s.	F.	Per cent. of significance
A	Sex	0·3	1	0·3	0·12	Nil.
	Age	67·55	7	9·65	4·02	0·1
	Residual	170·40	71	2·40	—	—
B	Sex	1·26	1	1·26	0·4	Nil.
	Age	97·49	7	13·93	4·11	0·1
	Residual	240·9	71	3·39	—	—
C	Sex	1·13	1	1·13	0·5	Nil.
	Age	73·98	7	10·57	5·08	0·1
	Residual	147·90	71	2·08	—	—
D	Sex	6·0	1	6·0	3·7	Nil.
	Age	77·1	7	11·01	6·79	0·1
	Residual	114·8	71	1·62	—	—
E	Sex	0	1	0	0	Nil.
	Age	50·75	7	7·25	9·69	0·1
	Residual	53·4	71	·75	—	—

No sex differences were discovered, but in all texts highly significant differences between age groups showed up.

We may not conclude that thought structures fall into certain age ranges but that formal thought occurs earlier in certain fields, such as Geography and Literature, than in others, such as History, in this test, and that this occurrence depends on the content of the field.

VII.—CONCLUSION OF STATISTICAL EVIDENCE.

To conclude the statistical evidence, and possibly to fill a gap in our knowledge of the mental development of the investigated groups, averages were obtained of intuitive, concrete and formal thought structures in terms of both Chronological and Mental Age

TABLE 5

AVERAGES. INTUITIVE, CONCRETE AND FORMAL THOUGHT ANSWERS PER CHILD IN CHRONOLOGICAL AGE GROUPS.

Group	C. Age	Intuitive Answers	Concrete	Formal	Number in Group
Junior	—	5·9	4·9	4·2	10
	11+	2·9	4·9	7·2	10
	12+	2·9	5·2	6·9	10
	13+	2·0	5·0	8·0	10
	14+	2·18	2·64	10·18	11
11+ to 14+	—	2·49	4·39	8·12	41
	15+	·64	2·27	12·09	11
	16+	·69	1·62	12·69	13
	17+	·93	1·27	12·8	15
15+ to 17+	—	·77	1·67	12·56	39

TABLE 6

AVERAGES. INTUITIVE, CONCRETE AND FORMAL THOUGHT ANSWERS PER CHILD IN MENTAL AGE GROUPS.

Group	Mental Age	Intuitive Answers	Concrete	Formal	Number in Group
Junior	—	5·42	5·58	4·0	12
	11+	3·58	4·67	6·75	12
	12+	3·0	4·0	8·0	5
	13+	2·82	4·45	7·73	11
	14+	0·75	3·5	10·75	4
11+ to 14+	—	2·87	4·34	7·78	32
	15+	0·55	1·67	12·78	9
	16+	0·83	1·83	12·33	6
	17+	0·77	2·0	12·33	31
15+ to 17+	—	0·74	1·91	12·35	46

Both in terms of Chronological Age and Mental Age, the figures show that from 13+to 15+ there is a considerable increase in the incidence of formal thought, with a corresponding decrease in concrete and intuitive level answers.

Answers at the concrete level, too, decrease with age but it is interesting to note that the averages of concrete level answers show little difference between the Junior group and the 11+ to 14+ group. The difference in formal thought in the two groups is balanced by the difference in intuitive level answers.

The close agreement between the three oldest groups may be partially caused by their comparatively narrow range of intelligences ; but it may, too, add support to Piaget's theory of equilibrium of formal operations at about 15 years of age. The incidence of formal thought before the age of 11 years would, however, tend to discredit Piaget's age limits.

VIII.—CONCLUSIONS.

(*a*) *Texts.*

Using the interview technique the texts, in order to be grasped by the testee, need to be short. This reduces their fruitfulness in terms of possible answers to questions seeking out formal thought, for they tend to lack material. Certain questions on texts produced answers which overlapped on previous answers. This was particularly in evidence in the answers to the questions on Text A and D. It would probably be better either to reduce the number of questions to two per text and increase the number of texts to nine, or to structure one's own texts to provide the ncessary material. Selection of the texts, and questions, necessarily involves considerable pilot work with a small, but representative, group of children in order to discover and circumvent such matters as general vocabulary difficulties and unreliable questions.

(*b*) *Questions.*

Even with careful pilot work it is difficult to obtain as homogeneous a group of questions on each test as one would like.

The results show that question 2, Text A, question 3, Text D, and question 2, Text E, are not homogeneous in comparison with their fellows.

(c) *Formal Thought.*

(i) It is clear that not all the children over a chronological or a mental age of fifteen have answered at the formal thought level all the questions set. However, almost all the chidlren in these age groups have scored thirteen or more out of fifteen answers at the formal thought level so that we may assume this to be equilibrium in Piagetian terms. On the evidence it can reasonably be stated that children with a chronological age of fifteen, of average or better intelligence will have reached equilibrium of formal thought.

(ii) All the individuals tested, regardless of chronological or mental age, scored at least one answer at the formal thought level. As few of the children tested were below the chronological age of eleven, and twelve were below a mental age of eleven, it would indicate the occurrence of formal thought well before the age of eleven stiuplated by Piaget (Table 7).

TABLE 7

Subject's Name	Chronological Age	Mental Age	Number of Formal Thought Answers
Marion, D.	7·8	9·1	6
Andrew, M.	8·10	8·4	5

(iii) Certain children of matching Chronological and Mental Ages had differing scores of formal thought (Table 8).

TABLE 8

Subject's Name	Chronological Age	Mental Age	Number of Formal Thought Answers
David, C.	11·9	11·4	12
Stewart, J.	11·9	11·4	5
Kay, D.	13·8	13·7	3
Margaret, C.	13·9	13·8	9

Overall, History Text B was least well answered in terms of formal thought. This may be accounted for by the inclusion in the text and questions of the concept of past and more distant past. This conclusion is supported by the work of Watts[1] who states that " the past perfect and future perfect remain purely literary for most children." He also asserts that an appreciation of the difference between the Perfect and Pluperfect tenses occurs last, if at all, in the understanding of tenses by school children. Dewey,[2] too, found that (i) the presentation of ideas in historical selections in sequence other than that in which the events occurred creates special difficulties for children, and (ii) most children possess a very rudimentary time sense, as shown by their difficulties in arranging events in proper sequence. With Watts we may conclude by saying that maturation, intelligence and experience each have a part to play in the development of verbal reasoning to the formal thought level.

(*d*) *Chronological or Mental Age Weighting.*

Carpenter[1] and Lodwick[6] found that total scores on Piaget-type tests were more closely related to Mental Age than Chronological Age. The results of the coefficients of correlation in this research show a similar weighting towards Mental Age, with the secondary modern the most representative sample, but this weighting is not significant.

Further, no coefficients of Reproducibility reached Guttman's criterion of 85 per cent. except those of the further education group in Texts C, D and E, and here there was no significant weighting towards either Chronological or Mental age. These results, and the results of the coefficients of correlation which were too low to be acceptable, confirm the previous conclusion on formal thought, i.e., the existence of another factor : experience, with maturation and intelligence in the development of verbal reasoning.

(*e*) *Sex.*

The analysis of variance shows no significant differences between the sexes.

(*f*) *Age.*

There are significant differences by age but they tend to be between blocks of ages rather than differences between one-year groups. Generally, too, there are significant differences between the 11+ and 12+ and the 15+, 16+ and 17+ groups. However, the analysis shows that the differences between the three groups is not so marked in the scores on the History texts, A and B, as it is in the scores on the Geography and Literature Texts C, and D and E.

(*g*) *Psychological Regression.*

The results support the conclusion that the incidence of psychological regression to the intuitive level is so high that it appears likely that answers at this level occur where experience, possibly verbal repertoire is lacking, for we find examples of regression occuring even amongst the further education groups, though their general answer level is that of formal thought.

The greatest regression, not unexpectedly, occurs in the answers to the History texts. Regression in the answers to the other texts, particularly in the further education group, seldom occurs.

IX.—General conclusions.

Formal thought in language subjects appears to need more than the attainment of a certain Chronological and/or Mental Age. This leads us to assume that other things such as cultural background, width of experience and verbal repertoire may be essential pre-requisites for the development of formal thought. The importance of the successful completion of,one stage of development in the development of the one to follow, as Professor Peel[7] points out, cannot be overlooked, and indeed should be a prime consideration of the educationalist. Where the necessary extra ingredients are lacking, we may expect regression to the earliest stages of pre-operational thought to occur.

Generally, the research tends to support Piaget's theory of structures of thought and his belief in the necessity of the earlier stages for the attainment of formal thought, but the incidence of formal thought in the Junior group makes one chary of accepting strict age divisions. From our own experience we know that well defined stages do not really exist. What is most important is the realisation that the reaching of equilibrium in one stage is essential for the development of the next. Therefore, the sequence of stages is perhaps more significant than their actual timing.

Proportional judgments appear from these tests to be attainable by children of 11 years or more, Chronological or Mental Age. It is interesting to note from the evidence that against the criteria C.A. and M.A., the percentage of formal thought follows a similar pattern. The younger age groups, 7—11, attain a percentage level of roughly 33·3 per cent. to 45 per cent. formal thought. By the age of 14 years, the percentage level has risen to the limits 60 per cent.—100 per cent. Figures show the line of demarcation to be in the region of 11 years C.A. or M.A. After the age of 14 years the level of formal thought is maintained.

ACKNOWLEDGEMENTS.—The writers are indebted to Dr. R. M. Beard for help in marking answers and to Professor E. A. Peel for help in planning this research.

X.—BIBLIOGRAPHY.

[1] CARPENTER, T. E. (1955). A pilot study for the quantitative investigaion of Jean Piaget's original work on concept formation. *Educ. Rev.*, VII pp. 142-149.

[2] DEWEY, J. C. (1900). Comprehension difficulties in history. *Research in the Three R's.* Hunnicutt and Iverson.

[3] FISHER, SIR RONALD A., and YATES, FRANK (1953). *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburgh, 4th Edition.

[4] GUTTMAN, L. (1949). A basis for scaling qualitative data. *Amer. Soc. Rev.*, Vol. 9, No. 2, April, 1944, pp. 139-150.

[5] INHELDER, B., and PIAGET, J. (1958). *The Growth of Logical Thinking.* Routledge and Kegan Paul, London.

[6] LODWICK, A. R. (1958). *An Investigation of the Question whether the Inferences that children draw in learning history correspond to the stages of mental development that Piaget postulates.* University of Birmingham, Diploma in Education dissertation (unpublished).

[7] PEEL, E. A. (1960). *The Pupil's Thinking.* Oldbourne, London.

[8] RAVEN, J. C. (Revised Order, 1956). *Standard Progressive Matrices. Sets A, B, C, D and E.* Univ. Press, Cambridge.

[9] TUKEY, J. W. (1949). Comparing individual means in the Analysis of variance. *Biometrics*, V, pp. 99-104.

[10] WATTS, A. F. (1944). *Language and Mental Development of the Child.* Harrap, London.

XI.—APPENDIX 1

(i) TEXTS AND QUESTIONS.

Text A. " The History of England." Arnold-Forster, p. 72. Dunstan.

He was a wise and honest adviser to those who asked for his advice and followed it. He had two great ambitions—the one to make England great and powerful; the other to make the Church rich and powerful. He was a stern, harsh man, and always anxious to have his own way; to those who disagreed with him he was rough and cruel. Such a man was sure to make enemies, but he did much that was good and is remembered as a great Englishman.

As Abbot of Glastonbury, he had learnt that many of the monastries were in bad order, and that priests and monks were leading lives that brought shame on the Church. He caused inquiry to be made, drove out the evil-doers, and put better men in their places. After his death, men called him a saint, and many strange stories were told of the wonderful things he was said to have done while he was still alive.

Question Number :

1.—Was Dunstan a good man? Why?
2.—Did he improve the Church of England or not? Why?
3.—Would you regard Dunstan as a great Englishman? Why, or Why not?

XII.—APPENDIX, 2.

(ii) SPECIMEN ANSWERS.

One answer at each of the three levels of thought is given for the first question only (Text A).

Text A. Question 1.

Was Dunstan a good man ? Why ?

Intuitive answer : Yes, he was. (Why ?)*.

Score one : Because he was a good man and did as he was told so that made him good.

Concrete answer . Yes, because he made the Church better (How ?).

Score two : He drove out the evil-doers.

Formal answer : He was in some ways, but not in others. (Why ?).

Score three : Because he was trying to make England and the Church powerful, but I don't think he was right in being cruel to those who disagreed with him.

* Questions in brackets are the supplementary questions put by interviewer.

Further examples of texts and specimen answers are available on request.

Manuscript received, 5th May, 1961.

THE ESTIMATION OF READABILITY OF BOOKS OF LOW DIFFICULTY

By JOHN McLEOD
(*Remedial Education Centre, University of Queensland*)

SUMMARY. The paper describes an experimental procedure designed to provide a meaningful index of the readability of four elementary reading books.

Two hundred-and-fifty-five children, whose reading ages had been assessed individually on a test of Word Recognition, were required to read thirty-word extracts from each book. From the proportions of children at different reading levels who read an extract with not more than two mistakes, each book's readability, expressed as a threshold in the manner of the Constant Method of psychophysics, was calculated. Reliable readability levels were obtained for each book.

A second, parallel experiment, in which children's reading ages were assessed by a reading comprehension test, was also carried out. A sample of 375 children was used. Readability levels which agree substantially with the earlier estimates, were again obtained.

Readability levels of the books according to the Spache and Dolch formulae bore no consistent relationship to the experimental findings nor to each other.

I.—PREVIOUS RESEARCH.

IN the course of our work at the Remedial Education Centre of the University of Queensland, it is important that backward readers should have access to reading material which they will find interesting but the difficulty of which is appropriate to their limited reading skills. An accurate estimate of the reading difficulty level of books is, therefore, very desirable, particularly at the more elementary stages. Generally, the difficulty of a new book has been estimated by subjective comparison with books which are more familiar. The purpose of the exploratory inquiry described in this paper was to investigate the feasibility of attempting an *empirical* estimate of the difficulty of reading books.

Previous research into the readability of books has been confined in the main to the United States and most of that has concerned itself with reading material at higher levels than those with which we are primarily concerned. Chall[1] has summarised some twenty-nine inquiries carried out over the past thirty-five years or so, but of these, nine only have dealt with reading material designed for use below grade III level; and of these only a couple—those by Dolch and Spache—dealt solely and reasonably adequately with the lower grades.

Dolch[2] sampled ten basal reading series by taking one sentence from each page of the books. Grade equivalents were related to median sentence length, the length of what Dolch termed the 'long sentence,' and to the percentage of 'uncommon words.' The mean of these three estimates was taken as the book's difficulty level. Dolch's conversion tables suggest that the method is not likely to be very discriminating, e.g., a 2 per cent. frequency of hard words places a book at Grade I level, while 4 per cent. raises it a whole year to the Grade II level. Dolch gives us no indication of the reliability or validity of his method and he accepted as his criteria of difficulty the publishers' grade designations.

Spache[4] also accepted publishers' grade placements as his criterion of book difficulty. He took 152 books commonly used in Grades I, II and III, from which he obtained 220 samples, each of 100 words. To pre-primers he assigned a grade level of 1·2, primers 1·5, first readers 1·8, second readers 2·1 and third readers 3·3. Two predictive variables were used: the average sentence length per 100 word sample, and a word difficulty measure. Analysis of his results produced a readability formula in the shape of a multiple regression equation. A multiple correlation of 0·818 was obtained between the two predictors and the criterion. This may appear impressive, but when it is remembered that the mean grade level of books used was 2·49, with a standard deviation of 0·86, it follows that his *predicted* grade level is subject to a standard error of 0·495. Thus, as Spache and Chall have each pointed out, formula predictions of reading difficulty levels are probably only accurate to within one full grade (or ± 1 year of reading age).

There is a vast and obvious difference between a book suitable for a reading age of 7 and for a reading age of 8, so that as sensitive measures of difficulty at lower levels, reading formulae are far from adequate, particularly in a non-American context. The word lists of Stone, Dale and Dolch, which are commonly used to assess the number of 'hard words' are American and perhaps obsolescent. Also, there is more than a possibility that readability formula may achieve an accuracy which is partly spurious, for it is not unlikely that modern elementary reading books are written 'according to formula' or at least with the research findings in mind.

A more fundamental criticism however, is that the foundation on which these formulae are based, i.e., the criterion, is often so flimsy. Frequenty, the rather cumbersome formulae appear to be little more than means of yielding approximate estimates of what in the first place was a fairly arbitrarily decided level of difficulty. And, granting for a moment that it may be possible to derive a formula that will accurately predict the reading difficulty level of a book as (say) reading age 7 years, what exactly does this mean? Does it mean that a child with a reading age of 7 years should be able to read the book fluently, or that he can manage it with a struggle, or what?

It was with the limited objective of trying to find out whether a meaningful index of the reading difficulty of a series of books could be established, that the present inquiry was undertaken.

II.—METHOD.

The method adopted was simply to take an adequate number of children each of whose reading ability was known and to find out whether they could read samples from a given book. The reading age at which 50 per cent. of children could read a sample was defined as the difficulty level of the book.

The books selected for examination were Books I, II, III and V in the Blue Series of "Wide Range Readers."* Experience with these books had indicated that they are a carefully graded series and each book is free from wide fluctuations of reading difficulty It was considered, therefore, that if coherent results were possible from the present experimental approach, they would be more likely to be achieved with a series such as "Wide Range."

A decision had to be made about what criterion of readability should be adopted—comprehension or word recognition. Comprehension may appear the more valid, but as Lorge[3] has pointed out, the method of measuring comprehension influences the rating of it. For instance, if questions are framed in unfamiliar words, fewer will be answered correctly, and any assessment of readability based on such questions will be too high.

---

*SCHONELL, F. J. and FLOWERDEW, PHYLLIS. *The Wide Range Readers.* Edinburgh· Oliver and Boyd

Over the range of reading ability being considered here, tests of word recognition such as those of Burt, Schonell and Vernon provide a trustworthy estimate of achievement, so in the first phase of the inquiry, word recognition was used as the criterion. In view of the success achieved however, a second experiment was later carried out using reading comprehension as criterion.

(i) *First experiment* (*Word recognition as criterion*).

Passages of thirty words, each commencing at the beginning of a sentence, were read from each page of the four books. Two-hundred and fifty-five children were tested, so that about a half of Books I, II and III was sampled: in the case of the larger Book V, the sampling was about 40 per cent. No *a priori* decision was made as to the number of reading mistakes to be taken as indicative that a passage was too difficult. As testing proceeded, however, it was judged that virtually all children making not more than two mistakes ina thirty-word passage were 'able to read' the passage. In the case of children who made three mistakes, the position was less clear, but it was felt that the majority was finding the task a little too difficult. Almost without exception, children making more than three mistakes were unable to read the passages at all satisfactorily. The criterion finally decided upon was that a passage was deemed 'readable' by a child if he or she made fewer than three mistakes in reading it. This criterion is at least definite and if it were to prove too harsh or too lenient, it could be amended and the experimental results re-analysed.

Each child's Reading Age was assessed individually by the Neale Analysis of Reading Ability test.* This test consists of a number of graded prose passages and involves the same sort of task as did the reading of a passage from the Wide Range readers.

(ii) *Experimental Procedure.*

Permission was obtained from the Queensland Department of Education to test children in a number of State schools. From the early results, it was clear that the books were too difficult for children with reading ages below 7 years and too easy for children of reading age higher than 10 years. Subsequent testing was, therefore limited, as far as possible to children within this range. Of the 255 children tested, twenty were in Grade II, 180 in Grade III and fifty-five in Grade IV. Mrs. Rorke, Senior Demonstrator at the Remedial Centre, tested sixty children, fifty-five were tested by experienced teachers attending the University course in Diagnostic Testing and Remedial Teaching and 140 by the writer. Most of the children came from four State schools, but altogether eleven schools were involved as well as a few individual children attending the Centre. No elaborately represertive sampling according to ability or socio-economic status was considered necessary, because any favourable or unfavourable circumstance which may have influenced children's ability to read passages from the books would be also reflected in their reading test scores.

Children were not tested in any particular order. Each child read one passage from each book, pages being taken in order, and reading age was assessed at the same test session. In half the cases, the Neale test was administered first; with the other half, the book passages were read first. Neale's standard instructions (apart from the reference to comprehension) were given before the child read the book passages as well as in the actual Neale test. The Reading Accuracy Age for each child and the number of mistakes in reading each of the passages were recorded. The whole procedure took, on average, 9 minutes for each child.

---

* NEALE, MARIE D. (1958). *Neale Analysis of Reading Ability*. London . Macmillan.

(iii) *Results.*

Children were grouped according to reading age at four-monthly intervals. For each group, the proportion achieving the criterion of reading a thirty-word passage with not more than two mistakes, was calculated. Distributions obtained were remarkably smooth, as can be observed in Table 1. The data were of the same form as in the psychophysical Constant Method and were, therefore, treated arithmetically by Spearman's summation method. (This is essentially the same as the process whereby mental age is calculated in a test such as the Terman-Merrill.)

TABLE 1

PROPORTION OF CHILDREN SUCCESSFULLY READING GIVEN PASSAGE, ACCORDING TO READING AGE.

Reading Ages (Neale : Accuracy)	Wide Range (Blue) Book			
	I	II	III	V
>10 : 0	1·00	1·00	1·00	1·00
9 : 8 to 9 · 11	1·00	1·00	1·00	0·86
9 . 4 to 9 7	1·00	1·00	0·85	1·00
9 : 0 to 9 : 3	1·00	1·00	1·00	0·80
8 · 8 to 8 11	1·00	0·96	0·79	0·79
8 : 4 to 8 7	0·96	0·74	0·57	0·66
8 : 0 to 8 : 3	0·78	0·70	0·48	0·52
7 : 8 to 7 · 11	0·72	0·40	0·47	0·25
7 . 4 to 7 · 7	0·27	0·24	0·11	0·05
7 . 0 to 7 : 3	0·25	0·25	0·11	0·07
6 : 8 to 6 . 11	0·18	0·12	0·06	0·06
<6 : 8	0·00	0·00	0·00	0·00

The reading difficulty levels of the books, and the standard deviations of the distributions, are recorded in Table 2.

TABLE 2

READABILITY OF WIDE RANGE (BLUE) BOOKS (WORD RECOGNITION).

Book	Threshold as Reading Age		S.D. of Distribution
	years	months	months
I	7	6·9	7·09
II	7	9·9	8·32
III	8	1·7	13·19
V	8	3·0	9·01

Being based on the effective results of some 200 to 220 children, an estimate could be made of the standard errors of the readability ages ; they were of the order of 0·6 of a month.

(iv) *Discussion of First Experiment.*

The primary aim of the experiment, that of establishing empirical estimates of the reading difficulty of the books, had been achieved. These empirical results were compared with assessments of difficulty arrived at by both the Dolch method and the Spache formula. (See Table 3.)

TABLE 3

COMPARATIVE ESTIMATES OF THE READABILITY OF WIDE RANGE (BLUE)BOOKS.

Book	Readability									
	Dolch Method		Spache Formula		Authors' Estimate	Experi-mental Results (Word Recognition)		Experi-mental Results (Compre-hension)	
	yrs.	mths.	yrs.	mths.	years	yrs.	mths.	yrs.	mths.
I	7	7	8	3	7—7½	7	7	7	11
II	7	6	8	9	7½—8	7	10	7	11
III	8	7	8	7	8—8½	8	2	8	2
V	9	7	10	1	9—10	8	3	8	7

In comparing the experimentally determined difficulty levels with the published estimates of Schonell and Flowerdew, it can be seen that there is a reasonable measure of agreement as far as Books I, II and III are concerned, but the experimental results suggest that the reading difficulty of Book V has hitherto been somewhat overestimated. On the other hand, it may be that for a book at this level, word recognition in itself does not adequately define reading difficulty.

(v) *Second Experiment* (*Comprehension as criterion*).

A second study was, therefore, carried out in which a similar experimental technique was employed, except that this time the criterion used was reading comprehension instead of word recognition.

Comprehension has the advantage that testing may be carried out on a group basis, but as was pointed out earlier, the wording of questions affects their difficulty ; further, comprehension may be interpreted in different ways.

In this instance, questions were constructed according to two principles

(1) They must be factual and literal, i.e., no interpretation, ability to draw inferences, etc., must be called for.

(2) The words used must not be more difficult than the actual text of the passage.

The questions were drawn up by two experienced remedial teachers, with the help of various common word lists. All questions were examined, and edited where necessary, by the writer.

(vi) *Experimental Procedure.*

Passages of not less than thirty words were chosen on each page of the books and enclosed within an ink framework. (The passages consisted of complete sentences.) Immediately underneath was fixed an appropriate comprehension question printed on white paper. The pages were cut from the reading books and reassembled as 128 four-page booklets, each containing one page from Wide Range Books I, II, III and V, respectively.

As a measure of reading comprehension, Schonell's Silent Reading Test A*, was administered to each child. Then the book passages were presented to the children, who had to read what was in the 'picture frame' and answer the question underneath, one minute being allowed for each question. Three-hundred-and-seventy-five children were tested in the course of this experiment

(vii) *Results.*

The results were tabulated, as shown in Table 4.

TABLE 4

PROPORTION OF CHILDREN CORRECTLY ANSWERING COMPREHENSION QUESTIONS FROM GIVEN BOOK.

Score on Schonell's Silent Reading Test A	Wide Range (Blue) Book			
	I	II	III	V
15—16	1·00	1·00	1·00	1·00
13—14	0·97	0·86	0·83	0·86
11—12	0·77	0·87	0·87	0·67
9—10	0·80	0·83	0·74	0·69
7— 8	0·74	0·60	0·59	0·41
5— 6	0·35	0·37	0·35	0·23
3— 4	0·24	0·32	0·18	0·08
1— 2	0·10	0·10	0·07	0·03
0	0·06	0·06	0·00	0·00

Once more using Spearman's summation method, the reading difficulty level of each book was calculated and recorded in Table 5.

TABLE 5

READABILITY OF WIDE RANGE (BLUE) BOOKS (COMPREHENSION).

Book	Threshold as Reading Age		S.D. of Distribution
	yrs.	mths.	months
I	7	11	15·17
III	7	11	16·19
III	8	2	15·02
V	8	7	13·61

The standard errors of the obtained readability ages are each of the order of 0·8 months.

---

* SCHONELL, F. J. and F. E. (1950). *Diagnostic and Attainment Testing.* Edinburgh: Oliver and Boyd.

(viii) *Discussion of Second Experiment.*

The readability ages calculated in the second experiment agree very closely with those obtained earlier. The difficulty of Book V is greater, relative to Book III, than in the first experiment, which is more in line with expectation. According to this second experiment, however, Book I appears to be comparatively difficult : this, it is believed, is due to the fact that no practice exercise was included. Despite the fact that standardised and comprehensive instructions were used, it was observed that some children found the situation somewhat unusual when they were called upon to answer the first exercise.

III.—General discussion.

The experiments described seem to indicate that an empirical approach to the assessment of book readability is feasible, but there are still difficulties to be overcome. On the the one hand, except perhaps for very easy books, the assessment of books according to a criterion of Word Recognition is open to question. Further, the administration of individual tests is rather time consuming, although if reliability comparable with that obtained from readability formulae will suffice, the testing necessary would not involve large numbers of children.

On the other hand, the use of comprehension questions introduces a variable of unknown magnitude. If questions are constructed which seem appropriate to the reading matter under examination, it is probable that the assessment of readability will be inflated spuriously in the case of more difficult books. If, however, the simple type of question used in this inquiry were to be employed, the readability of harder books would very likely be underestimated.

A technique which might well provide a satisfactory solution is Taylor's 'Cloze Procedure.'[5] In this method, a number of words are deleted at random from the passage being examined, subjects are required to fill in all blanks, by guessing from the context of remaining words what the missing words should be. This technique has already been used successfully at the teacher trainee level. If it can be successfully 'extended downwards' to children, its use in conjunction with the psychophysical treatment of data described in this paper, may well produce a sensitive, yet powerful combination for the empirical assesment of book readability.

IV.—References.

[1] Chall, Jeanne, S. (1958). *Readability*. Columbus : Ohio State University.
[2] Dolch, E. W. (1948). *Problems in Reading*. Champaign, Illinois : Garrard Press.
[3] Lorge, I. (1949). Readability formulae—an evaluation. *Elementary English*, 36.
[4] Spache, G. (1952-53). A new readability formula for primary grade reading materials. *Elem. Sch. J.*, 53, 410-413.
[5] Taylor Wilson L. (1953). Cloze procedure : A new tool for measuring readability. *Journalism Quarterly*, 415-433.
[6] Woodworth, R. S. (1951). *Experimental Psychology*. London : Methuen.

Manuscript received 7th *June*, 1961.

THE WECHSLER INTELLIGENCE SCALE FOR CHILDREN APPLIED TO A SAMPLE OF LONDON PRIMARY SCHOOL CHILDREN

By SHEILA JONES

(*Department of Child Psychology, University College, London*)*

SUMMARY. This study presents the results obtained on a stratified sample of London primary school children (forty boys and forty girls in each age group, at average ages 8, 9 and 10 years), tested with both the WISC and the Revised Stanford-Binet. The internal consistency reliability of the Full, Verbal and Performance Scales (and the individual subtests) of the WISC was determined for the three separate age groups, together with correlations of the WISC with the Revised Stanford-Binet and teachers' ratings of intelligence. A factor analysis of the sub-test inter-correlations at 8, 9 and 10 years is also included. Results for this London sample indicate :

(1) That for children of average and above average intelligence, a higher I.Q. is in general found on the Revised Stanford-Binet than on the WISC when the same child is tested with both. The size of the discrepancy between I.Qs. on the two tests is too large to permit the use of the WISC as a substitute for the Revised Stanford-Binet.

(2) The reliability coefficients for WISC Full, Verbal and Performance Scale are reasonably high for this sample, but the reliability of the individual sub-tests, especially for the younger children in the sample, is too low to warrant the diagnostic use of any but the most extreme deviations in sub-test scores.

(3) The WISC is more valid and reliable at year 10 than at year 8 or 9, for this sample. The factor analysis indicates that the Scale is measuring a common factor (provisionally identified with general intelligence) which accounts for approximately 31 per cent , 37 per cent., and 42 per cent. of the total variance at years 8, 9 and 10, respectively A second bipolar factor extracted at each age level, although barely significant, was tentatively identified as contrasting verbal and non-verbal abilities measured by the sub-tests, or more specifically, the tests requiring verbal comprehension with those depending on visual-spatial concepts.

I.—INTRODUCTION.

SINCE its introduction in 1949, the WISC has come into increasing use in this country as a supplementary individual intelligence test to the Revised Stanford-Binet. Before it can be accepted as an alternative to the latter, however, information is needed as to the applicability of American norms to British children, and on the reliability and validity of the scale and its individual sub-tests.

The obvious advantages of the WISC include a point scale with items of similar type graded in order of increasing difficulty, an equal number of verbal and performance sub-tests, and I.Qs. expressed in standard deviation units which obviates the difficulty of I.Q. variability at different age levels.

---

* Based on part of a thesis submitted for the degree of Ph.D. in the University of London (1959). The writer wishes to express thanks to Miss G. H. Keir, M.A., of the Child Psychology Department, and to the London primary school teachers whose help and co-operation enabled this research to be carried out.

Whatever the intrinsic merits of the WISC, however, its diagnostic value must largely depend on the extent to which the groups tested resemble those in the original standardisation sample, with regard to such characteristics as geographic, socio-economic and cultural environments. These last factors require special consideration when a test standardised in America is to be used with British children. Little published information has appeared about the use of the WISC in this country, but in 1952, the British Psychological Society set up a working party to " examine the value of the WISC, the extent of its use in this country as an additional individual intelligence scale and to consider what modifications, if any, need to be made in it for use with British children.'' The conclusions reached from a restricted pilot study on a sample of 200 normal children (aged between 10 years 4 months and 10 years 7 months) were that " the use of American standardised norms with British children is not justified . . . the need for restandardisation of the test for British children has been established . . . Further research upon the test is needed, particularly a factor analysis of the sub-tests and correlation studies with tests standardised in Britain.'' The results to be reported from this study are intended as a contribution to this urgently needed data, and are based on a stratified sample of 240 London primary school children, aged between 7½ and 10½ years, tested with both the WISC and the Revised Stanford-Binet.

II.—PROCEDURE.

The children were selected from fourteen primary schools in five boroughs of the County of London. Selection was based on age, sex and parental occupation, the latter being determined from data on social class distribution in London as a whole, as contained in the Census Report (1951) for London County. Schools were chosen to represent average, above average and below average types as judged by the socio-economic status of the pupils attending them and the areas in which they were situated.

All the children attended the junior departments of primary schools, which in London cater for children from 7 to 11 years. The children were divided into three age groups : (*a*) 40 boys and 40 girls aged 7 years 6 months to 8 years 5 months ; (*b*) 40 boys and 40 girls aged 8 years 6 months to 9 years 5 months ; (*c*) 40 boys and 40 girls aged 9 years 6 months to 10 years 5 months. Each child was tested individually at the school ; in every case the WISC was given first, followed by the Revised Stanford-Binet. The interval between the two tests varied between one and three weeks. (As the primary interest was in scores on the WISC, it was decided not to alternate the order of the two tests in order to avoid any possible contamination of the WISC scores. The Revised Stanford-Binet was thus liable to transfer effects as it was always given second.)

All the sub-tests of the WISC were administered in the order in which they appear in the WISC Manual (Wechsler, 1949), including the two supplementary tests (Digit Span and Mazes). Certain verbal modifications* were necessary before the WISC could be used with British children, and those used in this sample were essentially the same as those later recommended by the B.P.S. Working Party on the WISC (1955).

Teachers' ratings of the children were obtained on a five-point scale, ranging from ' well below average ' to ' well above average.' The teachers were asked to regard ' average ' not as the average level of their particular class, but rather as an absolute concept insofar as this was possible. They were also asked to assess the child on the basis of what he or she was capable of achieving, rather than on actual achievements.

---

* The modifications are reported in full in F. L. Young's M.A thesis (1953), University of London. Half of the sample reported on in the above study was tested by Mrs. Young and the other half by the writer.

III.—RESULTS.

A. *Mean I.Qs. and standard deviations.*

The mean I.Qs. and standard deviations obtained for the WISC are shown in Table 1, which also shows the corresponding values obtained for the Revised Stanford -Binet at each age level. All the obtained mean I.Qs. exceed 100, which for purely urban samples, such as this, is in accordance with other findings. For the Wechsler standardisation sample differences of 5·6 points between WISC Full I.Qs. for urban and rural groups were obtained (Seashore *et al* 1950). The mean I.Qs. for the urban group in Wechsler's total sample (1,327 children aged between 5½ and 15½ years) were 103·2 for WISC Full Scale (with a standard deviation of 12·9) 103·3 for WISC Verbal Scale (standard deviation 13·3) and 102·5 for WISC Performance Scale (standard deviation 13·5).

TABLE 1

THE MEAN AND S.D. OF I.QS. ON THE WISC AND R.S.B. BY AGE AND SEX.
(40 boys and 40 girls at each age.)

	WISC V		WISC P		FULL		Revised Stan-WISCford-Binet	
Age	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
8 years								
Boys	110·3	12·0	104·5	13·1	108·4	11·9	116·4	15·9
Girls	104·6	13·9	103·9	12·1	104·4	12·1	113·4	15·7
Combined ...	107·4	13·3	104·2	12·7	106·4	12·1	114·9	15·9
9 years :								
Boys	105·8	14·6	104·5	14·6	105·5	14·3	112·0	18·4
Girls	102·9	10·8	100·6	11·1	101·9	10·9	109·8	12·8
Combined	104·4	12·9	102·5	13·1	103·7	12·8	110·9	15·9
10 years :								
Boys	105·4	13·4	106·7	13·1	107·0	13·3	111·4	14·3
Girls	102·0	13·9	100·7	14·5	101·4	13·9	109·0	14·6
Combined . ..	103·7	13·8	103·7	14·1	104·2	13·9	110·2	14·5

The results for the Revised Stanford-Binet follow the trend for mean I.Qs. greater than 100 reported in other results for British children (Scottish Survey, 1949, Heasman 1955, and Marshak 1956). In the case of the youngest group in our sample, however, some bias may be present, as it is the custom in some London primary schools to delay entry from the infants' to the junior departments if the child is unable to read. The classes of 7 to 8 year old children may not all have been fully representative of this age group.

The standard deviations for both the WISC and Revised Stanford-Binet of the London sample are somewhat lower than those given for the standardisation samples at corresponding ages (Wechsler 1949 and Terman 1937), probably due to the homogeneous nature of our sample.

B. *Comparison of I.Qs. obtained with the WISC and the Revised Stanford-Binet.*

The mean I.Qs. obtained with the Revised Stanford-Binet at each age level in our sample were greater than the corresponding mean I.Qs. for all scales of the Wisc. The average difference is 7·2 I.Q. points and, as can be seen from Table 1, the widest discrepancies occur between scores on the Revised Stanford-Binet and the Wisc Performance Scale. Wisc Verbal I.Qs. are closer to those obtained for the Binet than either the Wisc Full or Performance I.Qs.

C. *Individual I.Qs. compared on the WISC and Binet.*

Individual differences ranged from 16 points higher on the Wisc full scale to 36 points higher on the Revised Stanford-Binet. Table 2 shows the percentage of children with given I.Q. differences between Binet and Wisc scores. 77 per cent. of the total sample obtained higher results on the Binet than on the Wisc Full Scale. The mean differences increase from ·07 points higher on the Binet for those with Revised Stanford-Binet I.Qs. of 90 to 99, to 5.5 points higher for those with I.Qs. of 100 to 109, and to 15.6 points higher on the Binet for I.Qs. of 130 and over. At each I.Q. level above 80, a greater proportion of children show higher ratings on the Binet than the Wisc. In general, it is at the lower levels of intellignnce that the scores obtained for the two tests are more likely to agree.

TABLE 2

Comparison of I.Qs. Obtained by Children Tested With the Wisc and The Binet.
Total sample=240 children.

Revised Stanford-Binet I.Q. Level	N	Higher on R.S.B.			Higher in WISC		No diff.	Mean diff. (R.S.B. -WISC)
		1—10 points	11—20 points	21 and more points	1—10 points	11—20 points		
130 and over ..	30	23·3	53·3	23·3	—	—	—	+15·60
120—129	39	35·9	51·3	7·7	—	—	5·1	+11·70
110—119	64	53·1	26·6	3·1	15·6	—	1·6	+7·10
100—109	52	57·7	17·3	1·9	15·4	1·9	5·8	+5·50
90—99	39	28·2	12·8	—	41·0	5·1	12·8	—0·07
80 and under .	16	27·5	21·4	—	50·0	—	—	+3·00

In using both the Wisc and the Binet it is usual in clinical practice to assign a category of intelligence corresponding to certain I.Q. levels. For instance, Wechsler provides seven different categories ranging from Mental Defective to Very Superior, based on the distribution of cases in a normal curve. In the case of the Revised Stanford-Binet applied to British children, it is customary to use the I.Q. categories agreed by the B.P.S. Committee of Professional Psychologists, which give nine categories ranging from Very Inferior to Exceptional Intelligence. It is of some importance to know whether a child judged ' Average ' on the Binet according to the B.P.S. classification would also receive a similar rating if tested with the Wisc according to Wechsler's classification, and similarly for other verbal description of categories. An analysis of each child's rating on the two tests is given in Table 3. Over half (134 children) of the

TABLE 3

SCATTERGRAM OF INTELLIGENCE RATINGS OBTAINED BY EACH CHILD ON THE WISC AND THE REVISED STANFORD-BINET.

N=240 children.

R.S.B. Ratings (B.P.S. Classification)	Ratings on the WISC (Wechsler's Classifications)						Total
	Very Superior	Superior	Bright Normal	Average	Dull Normal	Border-line	
Exceptional Intelligence	1	1	—	—	—	—	2
Superior Intelligence ..	6	11	11	—	—	—	28
Above Average	—	6	37	28	—	—	71
High Average	—	2	9	48	2	—	61
Average	—	—	—	39	5	—	44
Low Average	—	—	—	19	3	5	27
Limited Intelligence ..	—	—	—	—	5	2	7
Total	7	20	57	134	15	7	240

total sample obtained a rating of ' Average ' on the Wisc. The corresponding ratings obtained on the Revised Stanford-Binet for these children were as follows:

Thirty-nine children (or approximately one-third of the group) were also judged average on the Binet.

Forty-eight children (or just over a third of the group) were judged ' High Average ' on the Binet.

Twenty-eight children (about one-fifth) were judged ' Above Average ' on the Binet.

Nineteen children, the remainder of the 134 children who were rated ' Average ' on the Wisc, were rated ' Low Average ' on the Binet.

The only other group of appreciable size is that of thirty-seven children, judged ' Bright Normal ' on the Wisc and ' Above Average ' on the Binet. In general, the analysis in Table 3 indicates that children in the middle ranges of intelligence will be rated one category or more higher on the Binet than on the Wisc; 10 per cent. of our sample are children classified as ' Average ' on the Wisc, and ' Above Average ' on the Revised Stanford-Binet.

D. *Differences between scores on Verbal and Performance Scales of the WISC.*

The standardisation procedures for the Wisc were designed to result in a difference of zero points between the average of Verbal I.Qs. and the average Performance I.Qs. at each age level, and the actual differences obtained on Wechsler's samples differ only slightly from zero. Seashore's (1951) analysis of the standardisation sample for rural and urban groupings, however, showed that relatively more urban children have greater Verbal I.Qs. than Performance I.Qs., the reverse being true of rural children. The mean difference between Verbal and Performance I.Qs. for the urban group was about one point. In the present sample, the mean Verbal I.Qs. are greater than the mean Performance I.Qs. by 3·2 points at year 8, by 1·9 points at year 9 and by ·02 points at year 10.

Individual differences between Verbal and Performance I.Qs. for this sample were distributed as shown in Table 4. Half the sample show differences of 8 points or more, about a quarter show 15 points or more, and a tenth 20 points or more. These figures are similar to those obtained by Wechsler, which are also contained in Table 4 for comparison.

TABLE 4

Distribution of Verbal and Performance I.Q. Differences.

Size of V and P Differences (regardless of sign)	Wechsler's Standardisation Sample		This Study Sample	
	N	Cumulative Percentage	N	Cumulative Percentage
35+	14	100	2	100
30—34	36	99	4	99
25—29	68	98	9	98
20—24	140	95	15	95
15—19	284	88	24	87
14	68	75	3	77
13	74	72	6	76
12	83	69	11	74
11	88	65	8	69
10	82	61	8	65
9	117	57	16	63
8	112	52	15	56
7	163	47	20	50
6	114	40	14	41
5	131	34	14	35
4	132	28	17	30
3	124	22	16	22
2	144	17	13	16
1	140	10	17	10
0	86	4	8	3
Total N	2200	—	240	—

E. *Reliability.*

The internal consistency reliability of the Wisc sub-tests was determined for the present sample, using the Kuder-Richardson formula (1937). (No alternate form of the Wisc exists and it was not found practicable to retest the sample in the present case.) Table 5 shows the reliability coefficients for each sub-test at 8, 9 and 10 years, together with Wechsler's split-half reliability coefficients for ages 7½ and 10½ years (these being the only data available from the standardisation sample within the requisite age range).

TABLE 5

RELIABILITY AND STANDARD ERROR OF MEASUREMENT OF THE WISC TESTS.

Test							Wechsler's Sample			
	Age 8 yrs.		Age 9 yrs.		Age 10 yrs.		7½ yrs.		10½ yrs.	
N	80		80		80		100		100	
	r	SE_m	r	SE_m	r	SE_m	r	SE_m*	r	SE_m
Information	·58	1·81	·72	1·53	·73	1·46	·66	1·75	·80	1·34
Comprehension	·58	2·01	·64	1·80	·80	1·16	·59	1·92	·73	1·56
Arithmetic	·55	1·95	·56	1·79	·62	1·85	·63	1·82	·84	1·20
Similarities	·57	1·57	·64	1·80	·50	1·63	·66	1·75	·81	1·31
Vocabulary	·70	1·53	·70	1·42	·70	1·42	·77	1·44	·91	·90
Digit Span	·57	1·90	·70	1·59	·69	1·56	·60	2·45	·59	1·92
Picture Completion	·50	2·05	·60	1·77	·65	1·72	·59	1·92	·66	1·75
Picture Arrangement	·55	1·61	·61	1·56	·60	1·77	·72	1·59	·71	1·62
Block Design	·74	1·33	·68	1·58	·75	1·40	·84	1·20	·87	1·08
Object Assembly	·63	2·01	·43	2·19	·56	1·92	·63	1·82	·63	1·82
Mazes	·74	1·38	·59	1·54	·61	1·69	·79	1·37	·81	1·31
*Verbal Scale	·86	4·98	·87	4·65	·90	4·36	·88	5·19	·96	3·00
† Performance Scale	·80	5·66	·81	5·72	·85	5·46	·86	5·61	·89	4·98
‡ Full Scale	·89	4·02	·90	4·05	·94	3·40	·92	4·25	95	3·36

* Without Digit Span.
† Without Coding and Mazes.
‡ Without Digit Span, Coding and Mazes.

NOTES:

(1) Coding is a speed test and its reliability could not be estimated by this method.

(2) SE_m* is in Scaled Score Units for the subtests, and in I.Q. Units for the Verbal, Performance and Full Scales.

The reliability coefficients for Verbal, Performance and Full Scales are sufficiently high for most testing purposes, and as in the case of Wechsler's sample, the Verbal scale shows greater reliability than the Performance scale. The reliability of the individual subtests ranges from ·43 to ·80, with a tendency for values to increase with age level. The low reliabilities obtained for the 8 year level in our sample are close to those obtained for Wechsler's standardisation sample at age 7½ years. (Wechsler gives no reliability coefficients for ages below 7½ years in the Manual, but clearly these would be desirable.)

Consideration of the standard errors of the sub-test scores, which are also shown in Table 5, indicates that errors of measurement play a large part in deviations obtained on Wisc sub-test scores. The standard error indicates the amount of fluctuation that can be expected in sub-test scores, if the same subject were retested under similar conditions. The range of variation (at the ·01 level of confidence) is 2·58 standard errors on either side of the observed score. A child obtaining a score of 10 points on the Comprehension test at age 8 years, for instance, could conceivably vary on retest between a score of 5 points and a score of 15 points, due to imperfection of the measuring instrument.

F. Correlations between the WISC and the Revised Stanford-Binet.

Product-moment correlations between scores obtained on the Wisc and the Revised Stanford-Binet for the subjects in this sample are shown in Table 6.

TABLE 6

CORRELATIONS BETWEEN THE WISC, THE REVISED STANFORD-BINET AND TEACHERS' RATINGS.

(40 boys and 40 girls at each Age Level.)

Age Level	Revised Stanford-Binet *vs.*			Teachers' Ratings *vs.*			
	Wisc Verbal	Wisc Perf.	Wisc Full	Wisc Verbal	Wisc Perf.	Wisc Full	Revised Stanford-Binet
8 years .							
Boys	·77	·48	·72	·67	·43	·64	·58
Girls	·79	·46	·76	·73	·53	·75	·65
Both	·78	·47	·74	·70	·48	·70	·62
9 years :							
Boys	·89	·65	·90	·71	·57	·74	·76
Girls	·78	·58	·75	·72	·60	·72	·65
Both	·84	·61	·84	·71	·59	·73	·71
10 years ·							
Boys	·86	·64	·83	·74	·67	·77	·67
Girls	·90	·67	·86	·79	·57	·75	·78
Both	·88	·66	·85	·76	·62	·76	·73
Total Sample.... N=240	·84	·59	·81	·73	·57	·74	·69

For the total group of 240 children, correlations of ·84, ·59 and ·81 were obtained between the Revised Stanford-Binet and the Wisc Verbal, Performance and Full Scales, respectively. The Wisc is less verbal in content than the Binet, and this is reflected in the higher correlations obtained between the Revised Stanford-Binet and Wisc Verbal scores, compared with those obtained with both Wisc Full and Wisc Performance scores.

Table 6 also includes the correlation coefficients between teachers' ratings and the scores on the two tests. The average correlations of teachers' ratings with the Wisc Full Scale is ·74 and with the Revised Stanford-Binet is ·69. If the teachers' estimates can be taken as valid, then the Wisc appears to be as good a measure of general intelligence as the Binet.

G. Factorial validity.

The factor analysis of the sub-tests shown in Table 7 was carried out in order to determine whether the sub-tests are saturated with a common factor, and to indicate whether the common factor (if present) at one age level

TABLE 7

FACTOR ANALYSIS OF WISC SUB-TESTS.
(N = 80 at each age level)

	Year 8			Year 9			Year 10		
	Factors		Comm.	Factors		Comm.	Factors		Comm.
	I	II		I	II		I	II	
Information	·70	·16	·52	·79	·29	·71	·75	·32	·67
Comprehension ..	·53	·58	·61	·44	·52	·45	·61	·33	·46
Arithmetic	·58	−·22	·38	·58	−·16	·35	·67	−·06	·45
Similarities	·70	·20	·54	·67	·07	·46	·77	·02	·59
Vocabulary......	·71	·47	·72	·70	·41	·66	·81	·41	·81
Digit Span	·42	−·20	·22	·57	−·03	·33	·56	−·11	·33
Picture Completion	42	·06	·18	·45	·06	·21	·59	·17	·38
Picture Arrangement	·58	·09	·36	·52	−·25	·33	·64	·10	·42
Block Design	·66	−·41	·60	·75	−·31	·64	·73	−·17	·56
Object Assembly	·42	−·24	·23	·60	−·26	·43	·60	−·21	·41
Coding....... ..	·23	−·12	·07	·50	−·23	·31	·51	−·41	·44
Mazes	·54	−·39	·44	·68	−·12	·47	·39	−·33	·27
Contributory variance %....	31·2	9·2	—	37·7	7·0	—	41·6	6·5	—

corresponds to that at another. Burt's Simple Summation method was used to analyse the inter-subtest correlations at each age level (Table 8).

Examination of the first factor loadings in Table 7 shows that at each age level all the sub-tests are saturated with a common factor, the importance of which increases with age level, accounting for 31·2 per cent., 37·7 per cent., and 41·6 per cent. of the total test variance at years 8, 9 and 10, respectively. In comparing the magnitude of the common factor at each age, reference must be made to the reliability coefficients of the sub-tests (Table 5), which show a tendency to increase with age Tests with low reliability, coupled with possible specific variance, will show limited communality, and this will affect the average amount of common factor variance. The low reliability coefficients for year 8 probably accounts for the smaller first factor variance.

Consistently high first factor loadings are shown for the sub-tests Vocabulary, Similarities, Information and Block Design. The three 'alternate' tests, Digit Span, Coding and Mazes, on the other hand, do not have consistent first factor loadings, and tend to contribute less to the common variance. (These sub-tests were designated 'alternates' by Wechsler because of their relatively low correlations with the other sub-tests.) After extraction of the first factor, the residuals were tested for significance by comparing them with twice the standard error of the observed correlations of the sub-tests. For the 8 and 9 year groups just over 5 per cent. of the residuals were significant, and for the 10 year group slightly less than 5 per cent. were significant. A bipolar analysis was carried out on the residuals at each age level, and the second factor loadings are also shown in Table 7. The contrast between the sub-tests Information, Comprehension, Similarities, Vocabulary, all positively saturated, on the one hand, and Arithmetic, Digit Span, Block Design, Object Assembly, Coding and Mazes, all negatively saturated, is fairly consistent at each age level. Picture Completion and Picture Arrangement show generally small, almost negligible,

TABLE 8

WISC INTER-SUB-TEST CORRELATIONS FOR LONDON SAMPLE AT EACH AGE LEVEL.

Number=80 at each Age Level.

	YEAR 8										
SUB-TEST	1	2	3	4	5	6	7	8	9	10	11
1. Information											
2. Comprehension	45										
3. Arithmetic	48	21									
4. Similarities	62	44	39								
5. Vocabulary	60	63	32	57							
6. Digit Span	39	13	44	39	17						
7. Picture Completion	23	32	-04	25	34	07					
8. Picture Arrangement	35	37	33	44	46	06	30				
9. Block Design	34	13	36	37	34	30	28	31			
10. Object Assembly	16	05	22	19	09	15	33	33	47		
11. Coding	05	06	20	12	17	17	12	17	29	-01	
12. Mazes	32	02	38	24	21	25	26	31	50	49	10

	YEAR 9										
SUB-TEST	1	2	3	4	5	6	7	8	9	10	11
1. Information											
2. Comprehension	44										
3. Arithmetic	44	12									
4. Similarities	63	30	40								
5. Vocabulary	64	54	38	49							
6. Digit Span	49	16	52	40	47						
7. Picture Completion	37	29	21	26	35	18					
8. Picture Arrangement	35	14	33	30	21	17	27				
9. Block Design	46	18	42	54	42	34	27	56			
10. Object Assembly	32	18	21	39	23	30	37	47	62		
11. Coding	38	09	40	21	35	41	27	24	43	37	
12. Mazes	49	27	42	49	33	39	32	44	55	48	26

	YEAR 10										
SUB-TEST	1	2	3	4	5	6	7	8	9	10	11
1. Information											
2. Comprehension	56										
3. Arithmetic	55	38									
4. Similarities	56	52	64								
5. Vocabulary	73	64	52	61							
6. Digit Span	46	32	46	42	38						
7. Picture Completion	56	31	24	45	54	29					
8. Picture Arrangement	47	45	36	44	56	32	45				
9. Block Design	47	37	46	55	51	38	43	48			
10. Object Assembly	33	29	33	44	42	27	34	42	60		
11. Coding	16	19	38	37	27	35	28	31	41	45	
12. Mazes	23	13	32	24	16	31	25	19	32	30	27

positive saturations, with the exception of year 9 where Picture Arrangement has a negative saturation of the same order as the other negatively saturated tests. The dichotomy broadly suggests a division between verbal and non-verbal tests, or a contrast between manipulative and non-manipulative abilities.

IV.—DISCUSSION AND CONCLUSIONS.

The results reported above apply to a sample of London primary school children selected according to age and socio-economic representation, i.e., to a sample of relatively homogeneous educational background and purely urban in character.

(*a*) *Mean I.Qs. on the WISC and the Binet.*

For the 240 children between 7½ and 10½ years in this study, the mean I.Qs. were 104·8 on the WISC Full Scale, 105·2 on the WISC Verbal Scale, 103·5 on the WISC Performance Scale, and 112·0 on the Revised Stanford-Binet. The mean I.Qs. obtained on all scales of the WISC and for all age groups and both sexes are higher than the mean I.Qs. published for Wechsler's total standardisation sample. They are, however, comparable in value with those obtained for purely urban groups used in the standardisation sample. Krugman *et al* (1951), report on a sample of 332 New York children (selected to represent the various socio-economic groups), part of which was used for Wechsler's sample. Forty-four children aged 8½ years obtained a mean I.Q. of 104·73 on the WISC and 110·86 on the Binet; thirty-one children aged 9½ years obtained a mean I.Q. of 106·52 on the WISC and 116·84 on the Binet.

The mean I.Qs. obtained for this sample with the Revised Stanford-Binet, which was used as a control test, were also higher for all ages and both sexes than the values given by Terman (1937), but were comparable in size with values found for purely urban samples in America and Britain.

The mean I.Qs. on the Revised Stanford-Binet are higher for all ages and both sexes than the corresponding mean I.Qs. for all scales of the WISC. The WISC Verbal Scale I.Qs. were closer to those obtained for the Binet (mean difference 6·8 points) than either the WISC Full (mean difference 7·2 points) or Performance Scale I.Qs. (mean difference 8·5 points).

There was a tendency for greater differences between WISC and Binet I.Qs. to be associated with higher levels of intelligence (these differences favouring the Binet). Two-thirds of the children with I.Qs. of 120 and over as measured by the Binet were over 10 I.Q. points higher on this test than on the WISC, and a few of these differences were as high as 30 I.Q. points.

The differences obtained for our sample may have been an artifact of the experiment, since the WISC was always administered first. There may have been transfer from the WISC to the Binet. Several investigators (e.g., Pastovic and Guthrie 1951, and Krugman *et al* 1951) have alternated the order of testing, however, and have nevertheless obtained consistently higher I.Qs. on the Revised Stanford-Binet than on the WISC.

The size of the discrepancies between I.Qs. measured by the WISC and by the Revised Stanford-Binet indicates that the WISC cannot be regarded as a substitute for the Revised Stanford-Binet. Further research is necessary to investigate the consistent discrepancies between WISC and Stanford-Binet I.Qs. at different ages and different intellectual levels. These discrepancies may be due to underestimation of ability by the WISC or over-estimation by the Revised Stanford-Binet. Until more conclusive data are available, the American norms for the WISC appear, from present results, to be reasonably applicable to British children, provided caution is practised in their interpretation and in making comparisons with I.Qs. obtained by children measured with the Binet.

Comparison of intelligence classifications obtained by individual children on the WISC and Revised Stanford-Binet showed that the majority of children in the middle ranges of intelligence were placed in higher categories of intelligence by results on the Revised Stanford-Binet. Over 10 per cent. of the sample judged ' Average ' on the WISC were classified ' Above Average ' on the Binet.

(*b*) *Differences between Verbal and Performance I.Qs.* on the WISC varied considerably in magnitude for individual children. Half the sample showed differences of 7 points or more and a few of the differences were as high as 35 I.Q. points. Equal Verbal and Performance I.Qs. were the exception rather than the rule and differences should not be interpreted as necessarily having any diagnostic significance. Seashore (1951) has cautioned users of the WISC about interpretation of differences between Verbal and Performance I.Qs.: " A difference may be clinically important, but not just because it is a difference. Other data must be adduced to permit attaching any import to a discrepancy even as big as 5, 10 or 15 points."

(*c*) *Reliability.*

Using the Kuder-Richardson (1937) formula, the internal consistency reliability coefficients of the sub-tests were calculated at each age level. These ranged from ·50 to ·74 (median value ·58) for year 8; from ·43 to ·72 (median value ·64) to year 9; and from ·50 to ·80 (median value ·65) at year 10. To evaluate the magnitude of the reliability coefficients the standard errors of measurement were calculated. The estimated range of fluctuation of each individual's score (at the ·01 level) attributable to chance errors varied from between 5 to 10 points for the different sub-tests throughout the total age range.

The warning given by Seashore, *et al* (1950) against interpreting differences between sub-test scores, without reference to their reliabilities, is strongly reinforced by the values obtained in this study. A difference in score between, say, Vocabulary, and any of the other verbal sub-tests, would have to be at least five scaled score points before it could be said that any significant difference exists between the functions they measure, for any child in the present sample. The reliability of the sub-tests would appear to be too low to warrant the diagnostic use of any but the most extreme deviations.

The Full Scale reliability at the three age levels (8, 9 and 10 years) was ·89, ·90 and ·94, respectively. For corresponding ages, the Verbal Scale reliability was ·80, ·87 and ·90, and the Performance Scale reliability was ·80, ·81 and ·85 (i.e., lower than for the other two Scales).

(*d*) *Correlation with external criteria.*

Correlations between WISC Full Scale and the Revised Stanford-Binet ranged from ·64 to ·77, the average correlation beting ·74. The Verbal Scale correlates more highly with the Stanford-Binet than does the WISC Full Scale or the Performance Scale; the coefficients ranged between ·77 and ·90, the average value being ·84. The correlation between the WISC Performance Scale and the Stanford-Binet ranged between ·46 and ·67, the average value being ·59. It will be noted that for the youngest group (average age 8 years) the correlation between WISC Performance scores and the Stanford-Binet (Table 6) are much lower (·48 for boys, ·46 for girls) than any of the other coefficients at higher age levels. The Binet contains more items of the performance type at younger levels, and it might have been expected that WISC Performance scores would correlate less well with Binet scores as the latter scale became more verbal in content. The low correlations at year 8 may indicate that the WISC Performance items are more difficult at the 8 year level than the corresponding types of item on the Binet for this age range. The average correlation of the WISC Full, Verbal and Performance Scales with teachers' ratings were ·74, ·73 and ·57, respectively. The average correlation between the Revised Stanford-Binet and teachers' ratings was ·69.

Correlations obtained between the WISC and the external criteria are thus of sufficiently high order to indicate that the WISC is a valid measure of general intelligence, which covers the same general area of behaviour as the Binet, although the abilities measured by the two tests are not identical.

(e) *Factorial validity.*

In selecting items for inclusion in the WISC, one of the criteria used was that they should correlate well with other measures of intelligence. This, in turn, leads to the expectation that some degree of inter-item correlation will be present, i.e., that the items will measure a common factor. Wechsler (1958) has stated that the sub-tests are to be regarded as different measures of intelligence, not measures of different kinds of intelligence: "The grouping of sub-tests into Verbal and Performance Scales, while intending to emphasize a dichotomy as regards possible types of ability called for by the individual tests does not imply that these are the only abilities involved in the tests. Nor does it presume that there are different kinds of intelligence, e.g., verbal, manipulative, etc. It merely implies that there are different ways in which intelligence can manifest itself."

Factor analysis of the sub-test intercorrelations obtained for the present sample indicated that they were all measuring a common factor within each age group, although some of the sub-tests were definitely low in general factor saturation. For different age levels, the first factor saturations did not remain constant for a particular sub-test; there was a tendency for an increase in general factor saturation with age level.

The sub-tests which showed consistently high first factor saturation at each age level were Information, Vocabulary, Similarities and Block Design. The sub-tests, Digit Span, Coding and Mazes (designated as 'alternates' by Wechsler) showed generally low factor saturation at each age level. (Since the main function of the test as a whole is the measurement of the general factor, the inclusion of these alternative tests 'when time permits,' as Wechsler suggests, seems inadvisable.)

The common factor, which was provisionally identified with general intelligence, accounted for approximately 31 per cent., 37 per cent. and 42 per cent. of the total variance at years 8, 9 and 10, respectively. Differences in first factor variance were considered to be partly due to differences in reliability of the sub-tests at each age level.

A second bipolar factor extracted at each age level, although barely significant, was tentatively identified as contrasting the verbal and non-verbal abilities measured by the sub-tests. The line of division, however, did not correspond with Wechsler's classification of the tub-tests into Verbal and Performance Scales—the three Verbal sub-tests, Digit Span, Coding and Arithmetic, are grouped with the other Performance sub-tests possibly because they also involve visual-spatial qualities. The fact that Arithmetic shows a decrease in loadings with increase in age level is probably connected with the increase in verbal comprehension required in more difficult problems in this sub-test—the easier problems being confined to simple numerical relationships. This interpretation would lend support to the verbal *versus* non-verbal dichotomy already suggested.

Wechsler gives tables of inter-correlations between the sub-tests for years 7½, 10½ and 13½ in the WISC Manual (200 children at each age level) and Gault

(1954) factorized these tables for years 10½ and 13½, using Thurstone's centroid method, with a rotated bifactor solution. Four factors were extracted at each age level, a general eductive, spatial perceptual, memory and verbal comprehension factor. Two of the Verbal tests, Arithmetic and Digit Span, and two Performance tests, Picture Arrangement and Coding had the highest saturations on the memory factor. Maxwell (1959) also factorised Wechsler's data according to the method of minimum likelihood, omitting the sub-tests Information and Mazes from the matrices. Two factors, verbal-intellectual and space-performance, were extracted in each analysis, and as in our result, the grouping of sub-tests with highest loadings on these factors cut across the classification of the sub-tests into Verbal and Performance Scales.

The somewhat unstable factor pattern from one age level to another in our analysis may be simply attributable to the small numbers in each group (80 at each age level) which increases sampling errors. The performance sub-tests however, show less stability in factor pattern than do the verbal sub-tests. This may be due to the fact that five out of six of them are a mixture of speed and power, in that time bonuses are awarded for accurate performance. For the older groups, individual differences in test scores were more dependent on time bonuses as the majority of children were able to complete the tasks within the basic time limit. If the sub-tests measure speed and power in varying degree at different age levels, this may affect the factorial composition of the performance sub-tests from one age level to another.

V.—REFERENCES.

BRITISH PSYCHOLOGICAL SOCIETY—COMMITTEE OF PROFESSIONAL PSYCHOLOGISTS (MENTAL HEALTH). Final report of the Working Party on the WISC (1955).

BURT, C. (1940). *Factors of the Mind.* University of London Press.

GAULT, U. (1954). Factor patterns of the Wechsler Intelligence Scales. *Austral. J Psychol.*, 6, pp. 85-90.

HEASMAN, I. (1955). *Effect of Prematurity on Psychological and Physical Development.* Ph.D. Thesis, University of London.

H.M.S.O. (1954). *Census* (1951). Area report Greater London.

JONES, S. (1959). *A Statistical Study of the Wechsler Intelligence Scale for Children.* Ph.D. Thesis, University of London.

KRUGMAN, J. I., JUSTMAN, J., WRIGHTSTONE, W. J., and KRUGMAN, M. (1951). Pupil functioning on the Stanford-Binet and the Wechsler Intelligence Scale for children *J. Consult. Psychol.*, 15, 475-83.

KUDER, G. F., and RICHARDSON, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika,* 2.

MARSHAK, M. D. (1956). *A Study of the Wechsler-Bellevue Intelligence Scale applied to British Schoolchildren.* Ph.D. Thesis, University of London.

MAXWELL, A. E. (1959). A factor analysis of the Wechsler Intelligence Scale for Children *Brit. J. Educ. Psychol.*, 29, 237-41.

PASTOVIC, J. J., and GUTHRIE, G. M. (1951). Some evidence on the validity of the WISC. *J. Consult. Psychol.*, 15, 385-6.

SEASHORE, H., WESMAN, A., and DOPPELT, J. (1950). The standardisation of the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.*, 14, 99-110.

SEASHORE, H. (1951). Differences between verbal and performance I.Qs. on the Wechsler Intelligence Scale for Children. *J. Consult. Psychol.*, 15, 62-7.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1949). *The trend of Scottish Intelligence.* University of London Press.

TERMAN, L. M., and MERRILL, M. A. (1937). *Measuring Intelligence.* Harrap, London.

WECHSLER, D. (1949). *Wechsler Intelligence Scale for Children* (*Manual*). New York: The Psychological Corporation.

WECHSLER, D. (1958). *The Measurement and Appraisal of Adult Intelligence.* London: Baillière, Tundall and Cox.

YOUNG, F. L. (1953). *The Reliability and Validity of the WISC as Applied to British School children of Primary School Age.* M.A. Thesis, University of London.

(*Manuscript received, June 29th,* 1961.)

THE NEED FOR A SCHEMATIC LEARNING THEORY

By RICHARD R. SKEMP
(Department of Psychology, University of Manchester)

SUMMARY. Current learning theories have little direct application to problems of teaching and learning school subjects. The development of a schematic learning theory will, it is hoped, help to remedy this deficiency. Such a theory will also have wider applications.

I.—INTRODUCTION.

AMONG the branches of psychology applicable in the field of educational research, one would expect learning theory to be the most prominent of all. Learning is a child's chief task throughout his school days, and beyond: and the teacher's main concern is the extent and quality of this activity. How much help in these tasks have psychologists been able to offer, and teachers to use?

A search of the three main journals in the field of educational psychology and research, over the 21 years from 1940 to the year of writing (1961), has only revealed a total of five papers on applications of learning theory; the distribution being as follows:

British Journal of Educational Psychology—None.
Journal of Educational Psychology—Three.
Journal of Educational Research—Two.

This most important topic seems to be much neglected: and the reason is not far to seek. In one of the five papers listed above, Buswell* (1956) says: "Without the slightest criticism of experiments in general psychology, we cannot continue to be satisfied with implications for education from results of experiments with simple mental processes, with animals, and at the sub-language level . . . I am proposing that educational psychologists take their cues for research from the problems of learning in schools, where the processes are complex and where the learnings carried on are at the language level." In other words, theories relating to conditioning, reinforcement learning, sign learning, perceptual learning, etc., are not adequate for the classroom. A theory is required which takes account (among other things) of the systematic development of an organised body of knowledge, which not only integrates what has been learnt, but is a major factor in new learning: as when a knowledge of arithmetic makes possible the learning of algebra, and when this knowledge of algebra is subsequently used for the understanding of analytical geometry.

The only theory yet available which does this has been put forward by Piaget (1950). He calls such a body of knowledge a 'schema.' The incorporation of new knowledge into an existing schema is called 'assimilation'; and the enlargement of a schema, which may be necessary if it is not adequate for the above purpose in its existing form, is called 'accommodation.' These three related concepts would seem to offer a basis for the kind of learning theory which is needed.

---

* In his presidential address to the Educational Psychology division of the American Psychological Association, September, 1955.

Yet in a standard text on Theories of Learning, Hilgard (1958) makes no reference to it. The name of Piaget does not even appear at all in the list of authors referred to. This could mean either that Hilgard is not familiar with the work of Piaget, or that he does not regard it as applicable to learning theory. The latter hypothesis seems tenable, since the book by Piaget here referred to is called *The Psychology of Intelligence*; and Hilgard says (op. cit. p. 25) " . . . the problems of the nature of and measurement of intelligence lie outside the scope of this volume." This will perhaps seem strange to most teachers, who believe that intelligence and learning are rather closely related: and indeed, that the more intelligent a child, the better he can learn. But Hilgard does not thereafter refer to the subject of intelligence, and the learning theories he presents are indicated by the chapter headings: Thorndike's Connectionism, Guthrie's Contiguous Conditioning, Skinner's Operant Conditioning, Hull's Systematic Behaviour Theory, Tolman's Sign Learning, Classical Gestalt Theory, Lewin's Field Theory, Freud's Psychodynamics. This text is the only one on learning theory which is recommended to doctoral students by one-half or more of the American and Canadian psychology departments granting doctors' degrees, and appeared in this category in both 1953-4 and 1958-9 (Sundberg, 1960). If this text may on the foregoing evidence be taken as both representing and largely determining the state of learning theory in America, the lack of application of learning theory to education is easy to understand.*

In this country, however, the work of Piaget is rapidly becoming more widely known and appreciated, particularly during the last decade. Two authors who have played an important part in this are Peel (1956 and 1960) and Lovell (1960): who, long before the dates of these books, were (like the present writer) presenting Piaget's work to their students. Here, if anywhere, would one expect to find presented a schematic theory of learning.

In his earlier book, Peel discusses the following learning theories: association, S-R theory, conditioned responses, Hull's reinforcement theory and Tolman's theory of sign-learning; Gestalt psychology, and learning as insight; transfer. The index gives six references to Piaget, but in none of these is there any reference to a schema.

In the foreword of the second of the two works cited, Peel declares that " Piaget's work forms the core of the treatment." The word ' schema ' does not however appear in the glossary, nor in the index. I have found it twice in the text (Chapter VI), but not in relation to learning.

Lovell makes only two references to schemata. One is, however, important enough to be given in full, for several reasons. It points out that the concept of the schema was described earlier by Head and Bartlett, and was taken over by Piaget who developed it and added the concepts of assimilation and accommodation. Also it contains a clue why the idea has not yet been applied to learning theory.

> (page 97). "Much earlier, Head, the distinguished neurologist, had proposed that the cortex . . . was a storehouse of past impressions which normally remain outside consciousness. These impressions form organisations called *schemata* which modify all future impressions produced by incoming sensory data, so that all perception is affected by what has gone

---

*A colleague has since called my attention to a passage in the last chapter in which Hilgard himself points out the need for learning theories which are more relevant to real life situations, and refers to "the sterility of contemporary learning theories". I would therefore like to remedy a possible implication in my reference to his work, that he was less aware of their shortcomings than he here shows.

> before. Somewhat similar views on the building up and functioning of schemata have been proposed by Bartlett* and Hebb†. For Piaget *assimilation* and *accommodation* are fundamental processes. The former term implies the absorption and integration of new experiences into previously organised schemata, and the latter term the modification of schemata by new experiences or the build-up of new schemata."

This extract is in the chapter called " Imagery and Thinking," and it will be seen that the concept of the schema is considered only in relation to perception, and with a background of neurophysiology. In the chapter devoted to learning, it plays no part at all.

Without, therefore, doing injustice to the value of these books in providing an introduction to many aspects of Piaget's work, I think that it is fair to say that although there may be oblique reference to it, the concept of schematic learning as such does not appear in them. This is perhaps not surprising, for Piaget presents it in a book about intelligence, and the connection with learning is not made explicit : still less its implications for practical learning and teaching tasks. Again, no criticism is implied of this important but difficult work. It was not written as a text on educational psychology.

The purpose of the present paper is to put forward the view that only by developing further a theory along the lines which Piaget has begun can learning theory be made fully relevant for educational psychology (particularly since it offers a hope of relating certain kinds of learning to intelligence) ; and to describe an experiment which both supports this view and indicates the kind of way in which a beginning can be made with the practical application of the theory in the classroom.

II.—EXPERIMENT.

The chief object of this experiment was to demonstrate the crucial importance of a schema even for a relatively straightforward learning task, less difficult than most of those required of pupils in the usual school subjects.

To isolate the schema as the experimental variable, two different artificial schemata were devised, each based on sixteen basic symbols with associated meanings. These symbols were combined into pairs and threes, and then into successively larger groups ; symbolising new meanings which were not arbitrary, but based on the meanings of the sub-groups, which, in turn, were based on those of the individual symbols. This can best be understood by reference to Figure 1, which shows part of schema 1.

For experimental control of the other variables a second schema was prepared, similar in principle to the first. Twelve of the same sixteen basic symbols were used but with new meanings, together with four new ones. An extract from this second schema is shown in Figure 2. From considerations of space, it has only been possible to show a few of the more complex groups in these two figures. Apart from their content, however, it is hoped that they will help to clarify the concept by showing two examples of how schemata can be progressively built up.

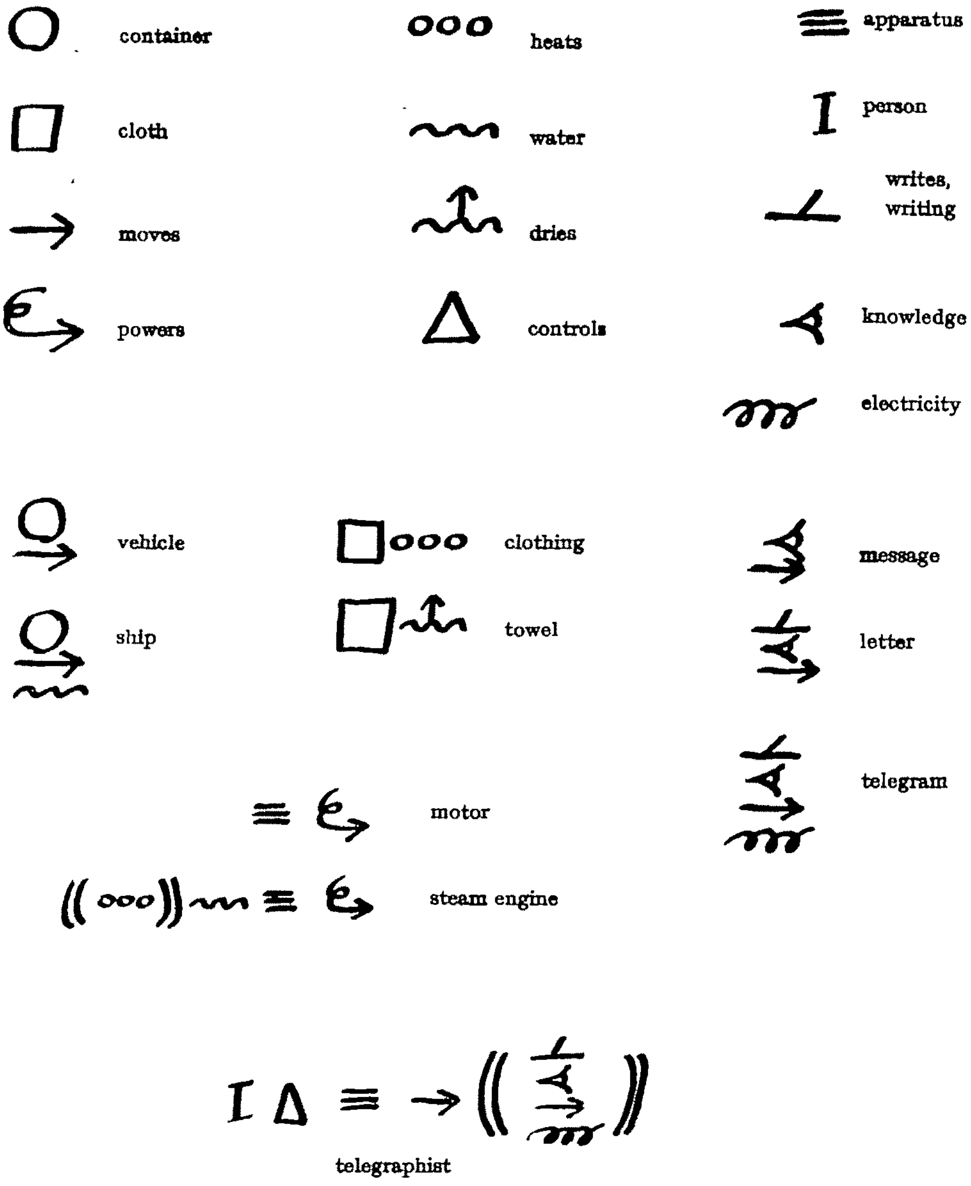
---

* Remembering.

† The Organisation of Behaviour.

PART OF SCHEMA 1

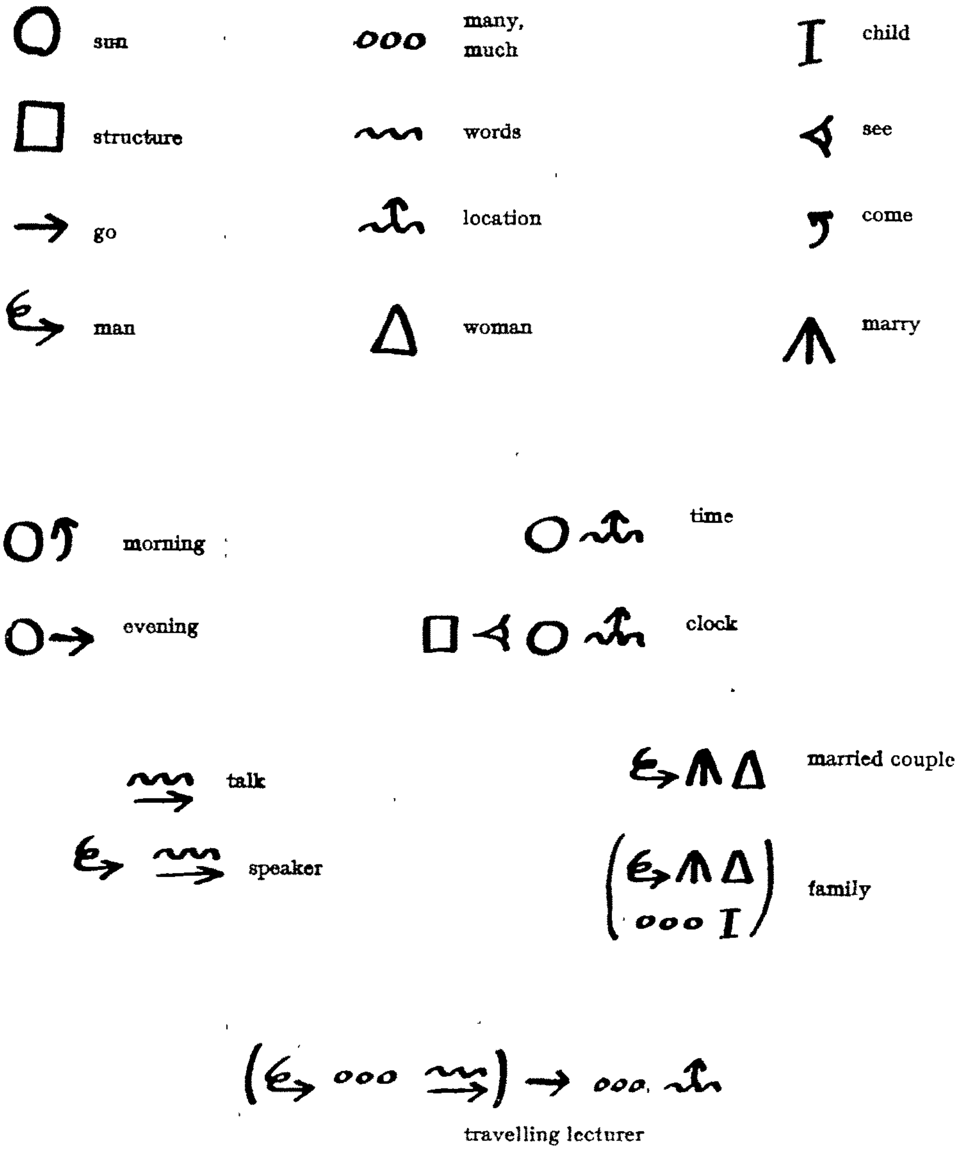
FIGURE 1



NOTE.— (( )) in this schema acts as a frequentative.

PART OF SCHEMA 2

FIGURE 2



The subjects were two parallel first forms at a boys' grammar school, numbering 23, 24, respectively, after excluding boys who were absent for any part of the experiment. In the preliminary part, each form learnt one of these schemata : that is, they were asked to memorise first the basic word and symbol pairs, and then the successively larger groups which symbolised new words. Pencil and paper were given to them, and they were advised that the best way to learn was by attempted recall with self-testing and correction of errors. This extended over three school periods on consecutive days, and the subjects were then ready for the final part of the experiment: which involved learning with, and without, the aid of the respective schemata which they had learnt.

For each schema a set of ten multi-symbol groups was prepared, these groups ranging from 8 to 14 symbols in size, and each set totalling exactly 100 symbols together. (One example from each final set has been included in Figures 1 and 2, being in each case the last one shown). Both forms were asked to learn both of these final sets : the task being, as before, to learn the group of symbols associated with each word. For one set, they had learnt a schema into which the new symbols and their meanings could be assimilated ; for the other, their own particular schema was no use. The aid given by the schemata could thus be quantified, by combining the scores which represented schematic learning for each group, and comparing these results with the combined scores on non-schematic learning (hereafter called rote learning). Both groups of subjects were learning the same total material, the two halves of which were approximately equal objective difficulty. There was, of course, a considerable subjective difference ; but this, the effect of the previously learnt schemata, is just what the experiment was concerned with.

TABLE I

SUMMARY OF EXPERIMENTAL DESIGN.

Subjects.	*Preliminary stage.*	*Final tasks and their nature for the subjects.*			
Group I	Learn schema I	Learn material assimilable to schema I	Schematic learning	Learn material assimilable to schema II	Rote learning
Group II	Learn schema II		Rote learning		Schematic learning

Composition of scores.

Schematic learning :	Group I, recall of material fitting schema I, combined with Group II, recall of material fitting schema II.
Rote learning :	Group I, recall of material fitting schema II, combined with Group II, recall of material fitting schema I.

The experimental design is summarised in Table 1. It took care of any slight difference between the abilities of the two forms, and also of whatever differences there may have been between the two final tasks. The number of presentations of the individual symbols previously to the final tasks were equal for the two schemata. It should further be emphasised that, in the final test, all the subjects were learning to associate new groups of symbols with new meanings. There was no question of learning different meanings for any of the basic symbols, or earlier groups of symbols, which they had already learnt.

III.—RESULTS.

In Table 2 are shown the percentages of symbols correctly recalled by all subjects immediately after learning, the following day, and twenty-eight days later (i.e., twenty-nine days after the original learning). By ' total scores ' is meant percentage recall of all the symbols in the final learning task. The ' controlled scores ' are calculated for those symbols only which occurred in both schemata, and which for each group constituted twelve of the sixteen originally learnt by them.

TABLE 2

PERCENTAGE RECALL BY ALL SUBJECTS (N=47).

	Immediate	1 day	28/29 days
Total Scores :			
Schematic learning	70	68	51
Rote learning	35	26	7
Controlled Scores :			
Schematic learning	69	69	58
Rote learning	32	23	8

In both cases, the difference between schematic and rote learning is striking. Schematic learning resulted in twice the number of symbols being recalled immediately after learning, and seven times as many after four weeks It gave both faster learning and better retention.

TABLE 3

AMOUNT FORGOTTEN (AS PERCENTAGE OF AMOUNT ORIGINALLY LEARNT).

	First Day	Following 28 Days
Controlled Scores ·		
Schematic learning	0	16
Rote learning	28	47

The rates of forgetting, expressed in percentages of what was originally learnt (as shown by immediate recall), are shown in Table 3. It can be seen that the subjects forgot more of the rote-learnt material in a day than of the schematically learnt material in four weeks.

IV.—IMPLICATIONS FOR TEACHING.

(*a*) *Schematic presentation of material.* The obvious consequence of these results is that, so far as possible, the subject matter of lessons should be arranged in such a form that pupils can use schematic learning rather than rote learning. What this means in its detailed application to the many subjects taught in schools and colleges has yet to be worked out. The main part of this task must of course, be the concern of the teachers who are specialists in various subjects :

but further psychological research is needed to establish more clearly what are the basic principles of schematic presentation. Probably the best teachers have always tended intuitively toward this method: but only by detailed investigation of the underlying processes can it be used most effectively.

(*b*) *Triple advantage of schematic learning*. The most obvious point established by the present experiment is the superiority of schematic learning over rote learning for both immediate recall and long-term retention. But two further advantages appear when the method is considered in the wider setting of school and college learning, and also in many of the learning situations of childhood and everyday life. In schematic learning, not only the material is learnt, but *the schema itself*: which is thus available to provide the same advantages for future learning (provided that this can be assimilated to the schema). Moreover, this subsequent learning, to the extent that it makes use of the schema, is also practising it: thereby providing, without extra work, revision of the earlier material. To give a simple example from algebra: a pupil who learns to factorise is also helping to prepare a schema for learning to solve many kinds of equation. When afterwards he is solving equations by this means, he is also revising factorisation. These advantages are absent from rote learning, in which successive tasks remain unrelated.

The same is true for some of the most basic schemata, such as speaking, reading and writing one's own language. Not only is later learning made possible (and also structured) by them, but they are constantly being practised and developed during the learning and exercise of more advanced topics. These basic schemata are, therefore, never forgotten, though not practised as such after the initial learning

Schematic learning thus has a triple effect: more efficient current learning, preparation for future learning, and automatic revision of past learning.

(*c*) *The basic schemata*. Another point arises in connection with these. When a child first enters an entirely new field of learning, elementary schemata have first to be built up *ab initio*. What other forms of learning may enter into this process is a matter for further investigation, but two points of importance for teaching seem clear. First: learning will be much slower at the stage when a new schema is being formed than later when it has become available for use as a tool of further learning. The teacher must, therefore, know at which of these stages his pupils are, and adjust his pace accordingly. Second: the experiment has shown how little is retained of material which cannot be assimilated to a schema; and the discussion of the previous section has indicated how self-perpetuating a schema tend to be. It therefore follows that whatever schemata are formed in the early stages of a child's education are of crucial importance for his future learning; and that any major changes of teaching method, beyond the powers of accommodation of existing schemata, could being progress almost to a standstill. One might well adapt William James' famous utterance about habit, and say that "we must guard against the development of schemata that are likely to be inappropriate for future learning as we should guard against the plague." Since the basic schemata are formed by children at an age before they are able to direct their own learning activities, it is parents and teachers who must guard them in this respect.

V.—SOME OTHER CONSIDERATIONS.

To consider all the implications for psychology in general of the concepts which have here been introduced, and the leads for further research which need to be followed to develop a useful body of knowledge about the processes of schematic learning, would be beyond the scope of this paper. Here I would like to mention two.

(*a*) *Learning theory and intelligence.* One of the criticisms here made of existing theories of learning, so far as they applied to educational psychology, has been their exclusion of the factor of intelligence. There may well be, even in adult humans, many learning processes into which intelligence does not enter. But where it is a factor at all, it is one which, to put it mildly, ought not to be to be ignored !

Piaget's concept of intelligence is a difficult one. " To define intelligence in terms of the progressive reversibility of the mobile structures which it forms is therefore to repeat, in different words, that intelligence constitutes the state of equilibrium towards which tend all the successive adaptations of a sensori-motor and cognitive nature, as well as all assimilatory and accommodatory interactions between the organism and the environment." (Op cit., page 11). It is hard to see immediately how this relates to the concepts of intelligence in the tradition of Binet, Spearman, Terman, Moray House, Wechsler, etc., which are familiar to British and American psychologists. Nor have we yet a Piagetian intelligence test, based on the above definitions.

Nevertheless, it is possible to see in the concept of a learning schema at least the possibility of a reciprocal assimilation (to use another concept from the same author) between theories of learning and of intelligence This follows from the intrinsic properties of concepts in classifying knowledge and inter-relating it in other ways also, thereby providing a ready means of schematisation. The formation and application of concepts has always been an important feature of traditional intelligence tests. Thus, the development of certain kinds of schemata can be regarded also as functioning of intelligence. Further development of our understanding along these lines would enable us not only to estimate the extent to which an individual is possessed of intelligence, but also to indicate how this intelligence might most effectively be used for various learning tasks. We could then, perhaps, teach children to use their intelligence intelligently.

(*b*) *Meaningful vs. non-meaningful learning.* Much work has already been done relating to this, by the Gestalt school and others. A major contribution has been that of Katona (1940), who showed with an abundance of experiment that meaningful learning was better retained than rote learning, and also gave better transfer. The reader might, therefore, reasonably ask what is new about the present study.

The crucial difference lies in the introduction of the schema, itself the result of learning, as that which gives meaning. Hitherto the emphasis has been on meaning as a property present or absent in the material itself. " We defined meaningful learning as learning which proceeds by organisation appropriate to *the inherent structure of the material.*" (Katona, op. cit., my italics). Where the meaning is not at once apparent, it is considered as resulting from a sudden perceptual re-structuring described as ' insight.' This gives an all-or-nothing character to the Gestalt approach. Insight either happens, or it does not, and the part played by learning in the process of organisation is not considered. An unfortunate result is that conversely, the part played by organisation in the process of learning has also been too little considered : in spite of its highly relevant subject matter, Katona's work is not mentioned by any of three books of Educational Psychology already cited.

The view here put forward is that meaning results from the existence *in the mind of the subject* of a suitable schema which, though it may be based on innate organisations, is chiefly the result of learning Meaning is thus synonymous with the process of assimilation to an existing schema: and by the development of an appropriate schema, meaning can be given to a set of hitherto unrelated data. Scientific theories are advanced schemata of this kind: and it is a far cry from this lengthy, systematic, and deliberate building up of meaning to the spontaneous perceptual restructuring of the Gestalt approach. Nevertheless, when the experiments of the latter are reconsidered in the light of the present approach, I think that they will be found to offer a useful contribution to a schematic theory of learning.

VI.—CONCLUSION.

In this paper an attempt has been made to clarify and develop the concept of a schema in the context of educational psychology; and to point out its central importance for school and college learning. Before concluding, however, it must be remembered that there are many other kinds of schemata besides the examples here given: for as here used, the concept includes all mental organisations which integrate existing knowledge and behaviour, determine its use in new situations, and form the basis for further learning. They certainly need not be conscious—the ability to reflect on one's schemata is a highly sophisticated one. They can be predominantly sensori-motor, such as driving a car; predominantly cognitive, such as mathematics; or have important elements of both, as in a surgical operation. Even the simplest trial and error learning Piaget believes to be not random, but directed by an existing and more basic schema: while, as already argued, complex cognitive skills such as mathematics and systems of science are also examples of schemata. Their varieties, development, internal organisation, transfer effects, and also their interaction with other factors of learning such as reinforcement and intelligence, are subjects for future research: having in every case applications within and beyond the field of educational psychology.

ACKNOWLEDGEMENTS.—I am most grateful to Mr. F. H. Philpot, Headmaster of Stockport Grammar School, to Mr. J. G. Gosling, and to the boys concerned, for their help in the experiment. Two student assistants, Mr. M. C. Kilcross and Mr. D. Castell, have also given valuable help in the preparation and scoring of the experimental material.

VII.—REFERENCES.

BUSWELL, G. T. (1956). Educational theory and the psychology of learning. *J. Educ. Psychol.*, 47, pp. 175-184.

HILGARD, E. R. (1958). *Theories of Learning* (second edition). New York: Appleton-Century-Crofts, and London: Methuen.

KATONA, G. (1940). *Organising and Memorising.* New York: Columbia University Press.

LOVELL, K. (1960). *Educational Psychology and Children* (fourth edition). University of London Press.

PEEL, E. A. (1956). *The Psychological Basis of Education.* London: Oliver and Boyd.

—— (1960). *The Pupil's Thinking.* London: Oldbourne.

PIAGET, J. (1950). *The Psychology of Intelligence* (English edition). London: Routledge and Kegan Paul.

SUNDBERG, N. D. (1960). Basic readings in psychology. *Bull. Brit. Psych. Soc.*, No. 42, pp. 47-50.

Manuscript received 18*th July*, 1961

SPECIFIC DYSLEXIA*

By M. D. VERNON

(*Department of Psychology, University of Reading*)

Summary. For the understanding of the nature and causes of backwardness in reading, it is important that backward readers should not be regarded as a homogeneous group whose disability merely varies quantitatively. It seems probable that not only do they differ in their responses to remedial treatment; but also that the nature of their reading difficulties, and the factors associated with these, differ in different cases. In particular, ' specific dyslexia ' seems to constitute a severe type of disability, particularly resistent to remedial treatment. Several investigations of this disability have recently been carried out, and evidence is adduced from them as to other associated symptoms, and to possible causes.

I.—Introduction.

In a previous paper (Vernon, 1960), the author noted that in spite of the numerous studies which have been made in recent years of reading backwardness, its incidence, causes and cure, there was still little reliable information as to the actual nature of this disability and its fundamental causes. This appeared to be due at least in part to the use of unsatisfactory techniques of investigation. In particular, the tendency has been to select and compare groups of normal and backward readers, and to attempt to isolate the factors which are associated with the latter group as a whole, but not with the former. Often the degree of backwardness has not been suitable measured; seldom has its exact nature been investigated in individual cases—that is to say, what exactly it is that each backward reader is unable to do or has failed to learn. The argument was put forward that we must study these children as individuals, and the peculiar nature of their disability; what were the particular pre-disposing, and possibly causal, factors in each case; and, if possible, their individual responses to remedial treatment. It was also noted that there seemed to be little recent contribution to the study of what might be a specific class of backwardness in reading, namely, ' specific dyslexia.'

Since the publication of this paper, further literature has been published which appears: (*a*) to reinforce the argument as to the erroneous conclusions which may be drawn if backward readers are studied only as a group; (*b*) to give us more information, albeit of a somewhat conflicting type, as to the nature and causes of ' specific dyslexia.' This new evidence suggests that there does indeed exist a particular type of reading disability which can be differentiated as ' specific dyslexia ' from other types of reading disability; that it appears relatively frequently among the more severe cases which are referred to clinics, but is undifferentiated and overlooked in studies of backward readers in schools.

II.—Classification of cases of reading backwardness.

That backward readers are not all of one type, at least in their responses to remedial reading teaching, may be illustrated by a consideration of a monograph recently published by Collins (1961). Here two groups each of twenty backward readers were given two types of remedial teaching, while a third control group continued to receive ordinary teaching in their schools. After six months' remedial teaching, the experimental groups had improved considerably on the average by comparison with the control group ; but follow-up testing at the end of 2½ years showed no difference between the average improvements of experimental and control groups. Collins deduced from this result that remedial teaching produces only a temporary spurt, and that in the long run, the children's reading progresses merely in pace with their normal rate of maturation. The remedial teaching is, therefore, ineffectual.

---

* Based on a paper read to the English Division of Professional Psychologists, in July, 1961.

Now, although this enquiry was admirably designed and administered, the conclusions drawn from its results must be considerably modified if the individual results are studied. It then appears that only 38 per cent. of all the children gained about as much in reading age as in chronological age, that is to say, 2½–3½ years : 36 per cent. gained more and 26 per cent. less. The children in the experimental groups seemed to improve slightly more than those in the control group, but the difference was not fully significant. It we study the children's final reading ages, we find that 58 per cent. of the experimental groups had reached a reading age of 10 years or over, as against 44 per cent. of the control group, the former being on the whole of lower intelligence than the latter. This probably means that these children could be said to read after a fashion ; at least they had mastered the basic reading processes. On the other hand, four children in the experimental groups and two in the control (about 13 per cent. of all those who reached the final test) were still semi-literate, with reading ages below 8 years. Two of these had I.Qs. below 90, but the other four were of normal intelligence.

It seems justifiable to conclude, therefore, that remedial teaching may be a help to some children, particularly those of somewhat sub-normal intelligence, and that they may be assisted by it to learn to read reasonably well. Others may be little affected by it, and go on at their own pace. A small minority learn almost nothing. It is surely, therefore, desirable in studying backward readers to discover if possible whether : (*a*) their difficulties are different in kind, as well as in amount ; (*b*) the causes, or at least the associated factors, are different; (*c*) certain types of remedial teaching are particularly effective in some cases and ineffective in others.

It should not be thought that a wide spread in the amount of improvement associated with remedial teaching is unique in Collins' investigation. Among recent studies, Hillman and Snowden (1960) showed that the improvement after three months in an adjustment class varied from eight months of reading age to over three years. Roberts (1960) studied the effects of three types of remedial teaching lasting twenty weeks, and found that 26 per cent. of the children had made less than six mouths' gain in reading age, whereas 33 per cent. had made more than twelve months. It is, of course, possible that these differences might have been smaller had follow-up testing been carried out at a later date. But we do not know what were the factors which produced the wide variations in immediate improvement. It does, therefore, seem important in future to study them and endeavour to understand them.

The author wishes to put forward the hypothesis, based more on general acquaintance with the literature relating to reading backwardness than on any particular investigation, that there are, in fact, three categories of backward reader :

(1) Those suffering from a moderate degree of backwardness, caused mainly by environmental factors such as lack of culture in the home, inefficient teaching, etc. These can often learn to read normally, or at least to grasp the basic phonic processes, if taught by suitable methods in small special groups.

(2) Those whose backwardness may or may not be severe, but who suffer from some form of emotional maladjustment due in the majority of cases to disturbed parental relationships, sometimes in conjunction with temperamental disorders. These usually require therapeutic treatment ; but if they respond to this, they may subsequently have little difficulty in learning to read.

(3) Those whose backwardness is often severe, and may be atrributable to some organic or constitutional factor, rather than to environmental factors. They are often of normal or even good intelligence, and may acquire proficiency in subjects which do not depend on reading. Emotional disorders, if they exist, appear to have been caused by the failure to learn to read. These are the cases which should properly be classified as ' specific dyslexics ' ; and it is they who form the main subject of this paper. They are often extremely resistant to remedial teaching. Even if they learn to read after a fashion, they may never do so fluently ; and they may never learn to spell correctly. Among the total school population their number may be relatively small, and they are usually overlooked in large scale surveys. But they form a considerable proportion of the cases referred to clinics for reading difficulties, and for this reason, they figure prominently in the reports of clinicians on reading disability.

Now it may often happen that these types overlap ; and that, for instance, mild cases of the third type become more severe because there are contributory environmental factors of the first type and/or some degree of emotional maladjustment Nevertheless, some relatively pure types do appear. It would, of course, be extremely interesting if it were possible to find out the particular stage at which each one broke down, and to discover corresponding differences in the processes which each found difficult. Unfortunately, in the present state of our knowledge, this is impossible. Nevertheless, it may be worth while outlining what we know of the difficulties encountered by backward readers. It seems possible that some of these may be particularly characteristic of dylexic cases.

III.—TYPES OF DIFFICULTY IN READING DISABILITY.

What then are the obstacles which arise in the progress of backward readers, and in particular what makes it difficult or impossible for them to acquire a grasp of the basic phonic processes ? Many of them may have learnt to recognize a few whole words the shapes of which they learnt by rote, by associating them with pictures. But this can hardly be said to constitute reading. Their failure to read might be due to inability to perceive and remember the shapes of printed letters, or to analyse word shapes into letter shapes They might be unable to associate the correct sounds with letter shapes ; or to combine the letter sounds into the sounds of whole words. From the author's observations of cases of reading disability, it appeared that failure in the earlier processes is comparatively uncommon. A more frequent cause of breakdown seems to be the lack of any *systematic* knowledge of phonic sounds and an inability to combine them together in the correct order. Thus, the author has examined a number of children who could pick out sounded words correctly from a group of printed alternatives ; and who could even build up sounded words from their separate printed letters. Yet they could not read the same words in the normal way. Either they were unable to synthesize separate sounds into whole word sounds ; or they could not synthesize them in the correct order. One might conclude that there was some deficiency in the manipulation of these sounds, and a failure to perceive the identity of the sequences of separate letter sounds with whole word sounds. Not only is it peculiarly difficult to break up the phonetic patterns of words into their constituent letter sounds; but equally, or perhaps more, difficult to identify a sequence of letter sounds with the phonetic pattern of the whole word. Thus, Hermann (1959) noted that dyslexics, whom he called 'word blind,' could not read words as entities, but tried to spell their way through them; and that words do not become entities by merely adding up their constituent letters.

That the organization and co-ordination of visual and auditory perceptions may well constitute a difficulty in learning to read seems very probable in the light of the experimental data and the hypotheses put forward by Piaget in *Les Mécanismes Perceptifs* (1961). He has demonstrated how primary perceptions are frequently deformed by 'centration' on certain parts of them, to the exclusion of other parts; and that 'perceptual activity' must supervene to correct inaccuracies by comparing and co-ordinating systematically all parts of the perceptal field. Now at the age at which children begin to learn to read, such perceptual activity is only beginning to develop, and is only feebly controlled by the 'schemes of action' which lead finally to completely logical reversible 'operations' in reading. Thus, it may be that at this age, the child is unable to think 'reversibly,' that is to say, to analyse the word shape into its letter shapes and re-synthesize their sounds to form the word sound, holding in mind what the word sound is like when it is broken up into letter sounds. It seems possible that dyslexic children remain particularly deficient in this capacity De Hirsch, for instance, in clinical studies of children with specific dyslexia (1954) discovered a general difficulty in the organization of patterns of both shapes and sounds, particularly of abstract shapes. Thus, these children were inferior in coding symbols in the WISC and in synthesizing the figures of the Bender test. In fact, they seemed to possess a general inferiority in the apprehension and matching of complex visual and auditory patterns.

It seems possible that backward readers are unusually deficient in the systematic appreciation of order and direction. There is no doubt that they are very liable to make mistakes over order and direction of letters in words; but it is not equally certain whether this is a general defect, or one which originates specifically in the all-round confusion of their reading processes. Ingram (1960) stated that there are two types of dyslexia, one of which is caused by difficulty in synthesizing separate sounds into meaningful units. These cases are characterized by retarded development of speech and of the capacity to understand speech in early childhood. In the other type there is no language disability, but there are 'spatial disabilities' in the recognition and reproduction of the correct orientation and direction of written symbols.

Many writers have supposed that there is some general failure in dyslexics to perceive and understand direction and orientation. Hermann (1959) noted that many 'word blind' individuals had remarkable difficulty in identifying right and left, and made mistakes over the discrimination between right and left. He compared the performance of 171 'word blind' cases, children and adults, with that of 295 primary and secondary school normal readers, on tests of ability to touch the left ear with the right hand, etc., and on finger naming. The 'word blind' made more mistakes than the normals in the first test (though 32 per cent. of the 'word blind' made no mistakes); and 29 per cent of the 'word blind' were uncertain on finger naming, as against only 9 per cent. of the normals. Reports of difficulties in left-right discrimination were given by about 70 per cent. of the 'word blind,' as against 25 per cent. of the normals Hermann thus considered that in 'word blindness' concepts of direction generally may be impaired, including the direction of letters in the visual images of words. This results from a disturbance of the 'body schema' related to impairment of function in the parietal lobes. However, we shall return to this hypothesis later.

Harris (1957) also found 'directional confusion,' as measured by difficulty and slowness in identifying the right hand, left ear and right eye, were much more frequent among seven-year-old backward readers (clinic cases) than among unselected school children of the same age. In Zangwill's study of twenty clinical cases of dyslexia (1960), aged 7–24 years, it was shown that confusion between left and right occurred in eight cases, and a tendency to scan with the eyes predominantly from left to right in five cases. Reversals in reading and/or writing were shown by six cases (all above the age at which these reversals normally disappear). Some cases showed repeated inability to copy the Necker cube figure correctly; and some fragmentation occurred in the copying of other drawings. Now some of these patients had several of these symptoms; some only one or two; and three showed no obvious symptoms other than the reading or spelling disability. Thus, 'directional confusion' is not a clear-cut entity, but may appear in several forms; and it is by no means universal in reading disability. This is apparent also from Ingram's and from Herman's cases.

IV.—THE SYNDROMES OF SPECIFIC DYSLEXIA AND THE FACTORS ASSOCIATED WITH THEM.

Many studies have produced evidence to show that other symptoms may be associated with the deficiencies just described. Among these are: slow speech development or speech disorders; incomplete or weak development of handedness (probably more significant than complete left-handedness or cross-laterality), hyperactivity and motor inco-ordination. Not all these symptoms occur in every case; indeed, it may be, as Ingram suggested, that cases with disorders of speech should be differentiated from those with 'spatial disabilities.' On the other hand, many of the symptoms may be found among other members of the family, suggesting a congenital basis. Thus, among Zangwill's twenty cases, only four showed no ambiguous laterality in either themselves or their families. Ingram found that young children with speech disorders, which frequently led later to reading difficulty, often possessed weak lateralization; and similar defects were noted in the patients' families.

The highest incidence of congenital factors in dyslexia appeared in Hallgren's study of Swedish children (1950); only thirteen of his 116 cases were said to show no incidence of dyslexia among relatives. Although his investigation of family histories was very detailed, the nature and extent of the actual reading disabilities were not clearly described. There was a high incidence of speech defects; but the investigation of laterality was unsatisfactory, and no conclusive evidence as to its frequency was obtained. Drew (1956) studied three related dyslexics, father and two sons, and found no signs of defective lateralization. He considered that in these cases some disability in the formation of Gestalten was inherited; and that in cases in which mixed dominance occurs, this may be secondary to a more basic disturbance. His cases were characterized by poor performance on the Wechsler Digit-Symbol test and by reversals in the Kohs Block test—the latter indicating the existence of 'directional confusion.' His evidence seems to agree, therefore, with that of Ingram and De Hirsch.

As to the neurological basis of these disorders there are several theories. Perhaps the best known is that of the failure to develop normal cerebral dominance. But other factors may supervene. Zangwill considered that cases of undeveloped dominance may be particularly vulnerable to various kinds of stress, including minor birth injuries of the brain, childhood illness and emotional maladjustment. Delacato (1959) also noted the frequency of severe childhood illness or head injury, associated with imperfect lateralization and ' directional confusion,' in 30–40 cases among 45 referred to a reading clinic for dyslexia. He postulated the existence of some form of brain damage less severe than that occurring in cerebral palsy. The similarity of these cases to cerebral palsied children is also suggested by De Hirsch's observations of the difficulties of dyslexics in figure-ground differentiation, and of fragmentation in reproducing the patterns of the Marble Board test.

Hermann postulated that ' word blindness ' resulted from some constitutional and congenital defect in the parietal lobe of the dominant hemisphere, a hypothesis also suggested by Drew. Hermann noted that the symptoms of Gerstmann's syndrome, caused by injury to the parietal lobe, are similar to those of ' word blindness,' and include ' directional confusion,' finger agnosia and impaired ability to write, to calculate and to copy drawings. Reversals in reading some times occur in Gerstmann's syndrome, but in general reading ability is only slightly imparied. This might be because these adult patients have reading habits too well established to be destroyed. But Eames (1960) stated that reading disability may occur in cases of Gerstmann's syndrome ; and that lack of development, injury or disease in area 39 of the parietal lobe may interfere with the ability to learn to read in childhood.

However, Ingram is quite clear that, in spite of the similarity of the symptoms, there are on the one hand, cases in which speech disorders and dyslexia are due to brain disease or injury acquired pre-natally or at the time of birth, which cases have no family history of similar disorders ; and on the other hand, cases in which there is no demonstrable abnormality of the brain, but with family histories of dyslexia, speech and motor disorders. It may, however, be practically difficult to differentiate these cases. Furthermore, although in many of the second type of cases laterality is weak, this is not so in every case. Thus, undeveloped cerebral dominance cannot be the only cause of congenital dyslexia.

Although the evidence is by no means conclusive, there appears to be reasonable support for the hypothesis that there does exist a number of cases who have great difficulty in learning to read, which difficulty seems not to be attributable to environmental factors or to emotional maladjustment. This difficulty may be accompanied in some cases by poor pattern or Gestalt perception (visual and/or auditory) and/or ' directional confusion,' weak laterality and impaired speech ; and similar characteristics may be apparent in other members of the family. The neurological basis or bases of these complexes are highly speculative ; failure in normal development of cerebral dominance, congenital defects of the parietal lobe and injury to the parietal lobe, have all been suggested. It is quite possible, as Ingram maintains, that dyslexia might be caused by different neurological deficiencies in different cases. But it seems clear that the exact nature of the disability cannot be diagnosed without much further and more accurate investigation ; nor, indeed, can we be certain at present whether only a single type of disability is involved, or more than one type. Clinical investigations are seldom carried out in sufficient detail, with properly controlled systematic testing. The existent tests are rarely very satisfactory, except perhaps Harris's tests of directionality and laterality (1956). We require suitable *ad hoc* tests of visual and auditory analysis and synthesis of complex forms ; and these tests must be standardized on suitable controls. Full data are needed, for every case, of lateralization and of perception and reproduction of direction. Full case histories are necessary, indicating possible brain injury, diseases and speech development; and evidence should be obtained as to the incidence of speech and reading disorders and of lateralization in other members of the family.

Since writing the above, the author has obtained a copy of *Right-Left Discrimination and Finger Localization* by A. L. Benton (New York: Hoeber and Harper, 1959), which describes very detailed and systematic investigations of several types of right-left discrimination. Benton found that some children showed a consistent tendency to right-left reversal, that is to say, that right-hand parts of their own bodies were systematically termed 'left,' and *vice versa.* This tendency was not associated with any lack of laterality, but was frequently associated with moderately or grossly retarded reading, and with considerable confusion over the transposition of left and right judgments to the bodies of others. These cases were contrasted with others who showed inconsistency in all left-right judgments and incomplete laterality. Benton attributed the defects of the former cases to some failure to develop accurate verbal formulations of left-right relationships, originating in language deficiencies. There seemed to be no evidence of cortical deficiencies such as lack of cerebral dominance or defects of the parietal lobe. It is clear that there is still much to be learnt about different types of 'directional confusion' and their relationship to reading, whether congenital or not. Benton points out that evidence as to the incidence of defects in right-left discrimination has often been vitiated by failure to explore these systematically.

V.—REMEDIAL TREATMENT.

Unfortunately, there is little or no reliable evidence as to the extent to which specific dyslexia, without marked emotional involvement, can be relieved or remedied by suitable treatment and teaching. In studies of the effects of remedial teaching no attempt is made to differentiate dyslexics from emotional and environmental cases. It is true that some studies, such as those of Fernald (1943), included children who were illiterate, or almost so, and who appeared to improve to a marked extent; but it is seldom clear in this or any other study of remedial teaching what was the exact nature of the disability or whether in the end they learnt to read normally. Indeed, very few long-term follow-up studies have been made. In reports on clinic cases, details of the effects of remedial teaching are rarely given. The general impression is that many at least of the more intelligent children learn to read after a fashion, particularly if given individual teaching; but they continue to read slowly and with difficulty, and to spell incorrectly. Thus, it has often been assumed that they are suffering from an ineradicable defect. Possibly their recovery depends as much on interest and motivation as upon the particular method of teaching them. However, these are little more than speculations. We need to know far more about the nature of specific dyslexia and about the effects on it of different types of teaching before we can arrive at any final conclusions.

VI.—REFERENCES.

COLLINS, J. E. (1961). The effects of remedial education. *Educational Monographs, University of Birmingham Institute of Education, No.* IV.

DE HIRSCH, K. (1954). Gestalt psychology as applied to language disturbances. *J. Nerv. Ment. Dis.*, 120, 257.

DELACATO, C. H. (1959). *The Treatment and Prevention of Reading Problems.* Springfield, Ill.: Thomas.

DREW, A. L. (1956). A neurological appraisal of familial congenital word-blindness. *Brain*, 76, 440.

EAMES, T. H. (1960). Some neural and glandular bases of learning. *J. Educ.* 42, No. 4.,

FERNALD, G. M. (1943). *Remedial Techniques in Basic School Subjects.* New York: McGraw-Hill.

HALLGREN, B. (1950). Specific dyslexia. *Acta Psychiat. Neurol. Suppl.*, 65.

HARRIS A. J. (1956). *How to Increase Reading Ability*, 3rd ed. New York: Longmans Green.

HARRIS, A. J. (1957). Lateral dominance, directional confusion and reading disability. *J. Psychol.*, 44, 283.

HERMANN, K. (1959). *Reading Disability.* Copenhagen: Munksgaard.

HILLMAN, H. H., and SNOWDON, R. L. (1960). Part-time classes for backward readers. *Brit. J. Educ. Psychol.*, 30, 168.

INGRAM, T. T. S. (1960). Paediatric aspects of specific developmental dysphasia, dyslexia and dysgraphia. *Cerebral Palsy Bull.*, 2, 254.

PIAGET, J. (1961). *Les Mécanismes Perceptifs.* Paris: Presses Universitaires de France.

ROBERTS, G. R. (1960). A study of motivation in remedial reading. *Brit. J. Educ. Psyhcol.*, 30, 176.

VERNON, M. D. (1960). The investigation of reading problems to-day. *Brit. J. Educ. Psychol.*, 30, 146.

ZANGWILL, O. L. (1960). *Cerebral Dominance and its Relation to Psychological Function.* Edinburgh: Oliver and Boyd.

Manuscript received, September 27th, 1961.

INFLUENCE OF TOWN AND COUNTRY UPON CHILDREN'S INTELLIGENCE

By HARBEN BOUTOURLINE YOUNG and RENATO TAGIURI
(*Harvard University*)

GINO TESI and GABRIELLA MONTEMAGNI
(*Harvard-Florence Research Project*)

SUMMARY. The results of a relatively culture-free intelligence test applied to a random sample of fifth grade elementary school classes in the Florence (Italy) area showed a marked superiority of town boys over boys from the country. This superiority remained significant for socio-economic classes 4 and 5 after differences in social class distribution had been controlled. However, no differences between urban and rural girls were found.

Possible explanations of the results are discussed and the importance of further enquiry is emphasised.

I.—INTRODUCTION.

As part of the activities of the Harvard-Florence Research Project* Raven's Progressive Matrices (1956), a relatively 'culture free' intelligence test, was applied to a sample of classes from the elementary school districts that feed into the junior high school system (6th to 8th grades) of Florence, Italy.†

The test results showed that boys in the *city* of Florence obtained higher scores than girls (mean percentiles: boys 59·06, girls 49·39), while in neighbouring *rural* areas the situation was reversed, girls scoring higher than boys (mean percentiles: boys 42·03, girls 49·71). Town and country girls, however, performed equally well.

The differential nature of the results for boys and girls indicates that this is not one more instance of the common observation that on intelligence tests city children do better than rural children (e.g., Anastasi, 1958, 1960). It is, indeed, possible that these findings may be manifestations of factors not yet understood, yet of potential consequence to the educator. In this paper we consider some of the ways in which such results could be explained.

II.—SUBJECTS AND PROCEDURES.

The data on elementary schools in the area and their fifth grade populations were supplied by the Ministry of Education. A 15 per cent. random sample of fifth grade classes was then chosen from town and country (town, 793 children in 38 classes; country, 373 children in 20 classes) from a total number of 7,959 children. As far as is known, none of the children used as subjects had previously undergone intelligence testing.

---

* The Harvard-Florence Research Project is a research group studying relationships between physical and mental growth and socio-economic conditions and is particularly interested in bodily and mental changes induced by environmental pressures. It is sponsored by Harvard University School of Public Health and is associated with the Universities of Florence, Rome and Palermo.

† This was done in the course of investigating the relative importance of intelligence, socio-economic conditions, sex, town or country dwelling, ease of travel to secondary school and other factors as to whether or not the child will go from the elementary to the classical junior high school. This study was necessary to elucidate factors of school selection for a cross-cultural study which draws children from the school systems of the United States and Italy. A description of the relative importance of these factors will be made in a later report.

The mean age of all the children studied was 11 years 2 months, with a range from 9 years 6 months to 14 years 6 months. For the city children the mean was 11 years 1 month, with range from 9 years 6 months to 14 years 6 months. In the country the mean was 11 years 3 months, ranging from 9 years 6 months to 14 years 6 months. Table 1 shows the distribution of the subjects according to age, sex and city or country.

TABLE 1

DISTRIBUTION OF SUBJECTS, ACCORDING TO AGE, SEX AND CITY OR COUNTRY.

Age	City		Country	
	Males %	Females %	Males %	Females %
9½	1	—	1	1
10	3	2	—	1
10½	14	14	12	16
11	20	18	17	15
11½	5	6	8	7
12	3	4	5	5
12½	2	2	5	2
13	2	2	1	2
13½	1	1	—	—
14 or over	—	—	1	1
	51	49	50	50
	793 Subjects		373 Subjects	

Raven's Progressive Matrices were administered by personnel of the Harvard-Florence Research Project. One hour was allowed for the test. The proportion of children not completing the test within one hour was 4·6 per cent. The mean time for completion was 39 minutes. In town, the boys took an average of 36 minutes 12 seconds, the girls 39 minutes 07 seconds. In the country, the average time taken was 41 minutes 48 seconds for the boys and 41 minutes 46 seconds for girls. Each child also gave the parental occupations and address, and stated whether he intended to proceed further with education and, if so, which alternative he would choose. There are four possibilities open in this latter respect: (*a*) to proceed to classical junior high school, (*b*) to proceed to technical junior high school, (*c*) to continue with 6th and 7th grades, which exist in some elementary schools, (*d*) to discontinue school.

III.—RESULTS AND FURTHER INVESTIGATIONS.

Table 2 shows the distribution of results on the intelligence test. The scores are grouped into three categories, 0–33, 34–66, 67–100 percentile ranges, respectively, and correspond to poor, average and good performance on the Raven. Males show a concentration of high scorers in the town and of low scorers in the country. Females show no difference in distribution between town and country. The chi-square ratio (d.f.2) is 17·0 for the town and 16·3 for the country, each with a probability of less than 0·01. In the remaining part of the paper we attempt to account for these results.

TABLE 2

PERCENTAGE OF SUBJECTS IN EACH CATEGORY OF SCORES.

Test Scores (percentiles)		0–33	34–66	67–100
City	Males	13	14	26
	Females	16	14	17
		29	28	43
Country	Males	25	14	8
	Females .	17	17	19
		42	31	27

One possibility is that there exists a process of selection that in the town tends to reduce the number of the less gifted among the males, while in the country more of the gifted leave school. There is, however, no evidence for this ; the percentage of males and females is practically equal grade by grade in the town and country and is in line with what would be expected from examination of the birth rates and infantile and child mortality figures for these age groups. There is no significant age difference between the town and country groups.

Since the performance of the girls is equal in town and country, such influences as local migration toward town, because of superior family intelligence, seem unlikely.

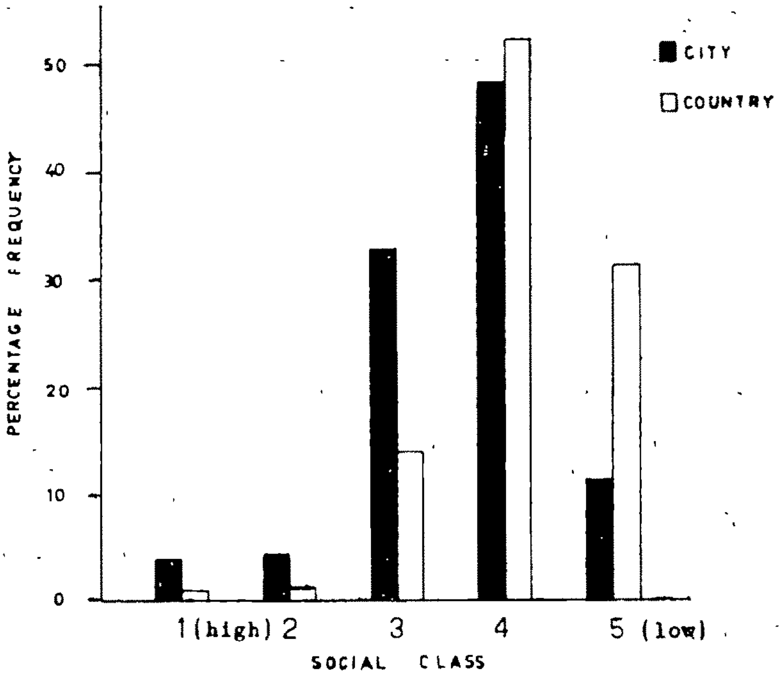
The effect of mixed classes should be considered as another possibility In the city it is unusual to find mixed classes in the elementary school. In the country approximately half of the classes are mixed. Results of males and females from country schools with separate classes were compared with results from mixed classes. It was found that, if anything, boys in mixed classes performed better than boys in all male classes. The factor of mixed classes, then, does not appear to account for the results.

Socio-economic levels should be considered as another possible basis for the difference observed. Figure 1 shows the distribution of the children into five socio-economic groups, from 5 (low) to 1 (high), derived from parental occupation and neighbourhood of dwelling. Town and country children can be compared in terms of percentage of total population in each group. The percentage of town children in social class 3 is twice as large as that of rural children. Class 4 is closely comparable and the inverse situation to class 3 is seen in class 5 where country children are more than twice as frequent as city children. This is due to the very large number of agricultural workers (class 5 occupation).

There is a known relationship between social class and intelligence (Havighurst and Janke 1944, Havighurst and Breese 1947, Davis 1948). Figure 2 and Table 3 show that this relationship holds for our subjects. Intelligence increases with socio-economic level. This is true for both sexes (males and females considered together show a progressive and significant increase in performance from class 5 through class 3, both for town and country). The performance of the girls within each of these social classes is similar in town and

FIG. 1

DISTRIBUTION OF CITY AND RURAL SUBJECTS AMONG FIVE SOCIAL CLASSES.



country, while for the boys there are marked town-country differences (class 4, p < ·001, class 5, p < ·02). Table 3 shows that in the town there is a significant superiority of boys over girls from social classes 4 through 2. In the country there is a constant superiority of girls over boys from social classes 5 through 3, but these differences are not significant.

It may be concluded that the differences in social class distribution do not satisfactorily explain the results seen in Table 2.

The changes brought about by the onset of puberty should be considered as another possible explanation. At the age of 11 years some 40 per cent of the girls in the Florence area will be in the first stage of puberty* and this might possibly affect performance. There are likely to be more early-maturers (Tanner 1955) in town, where socio-economic conditions are generally better, but it has already been noted that there are no significant differences between the performance of girls in town and country. The mean age for the commencement of puberty in boys for the Florence area is just under 13 years.* The proportion of town boys 12½ years or more was 11 per cent.; that of country boys was 15 per cent. Therefore, the physical and emotional changes of puberty are not regarded as a likely explanation.

* Harvard-Florence Research Project, unpublished data.

FIG. 2

PERFORMANCE ON INTELLIGENCE TEST AND SOCIAL CLASS DIFFERENCES BETWEEN MALES AND FEMALES.

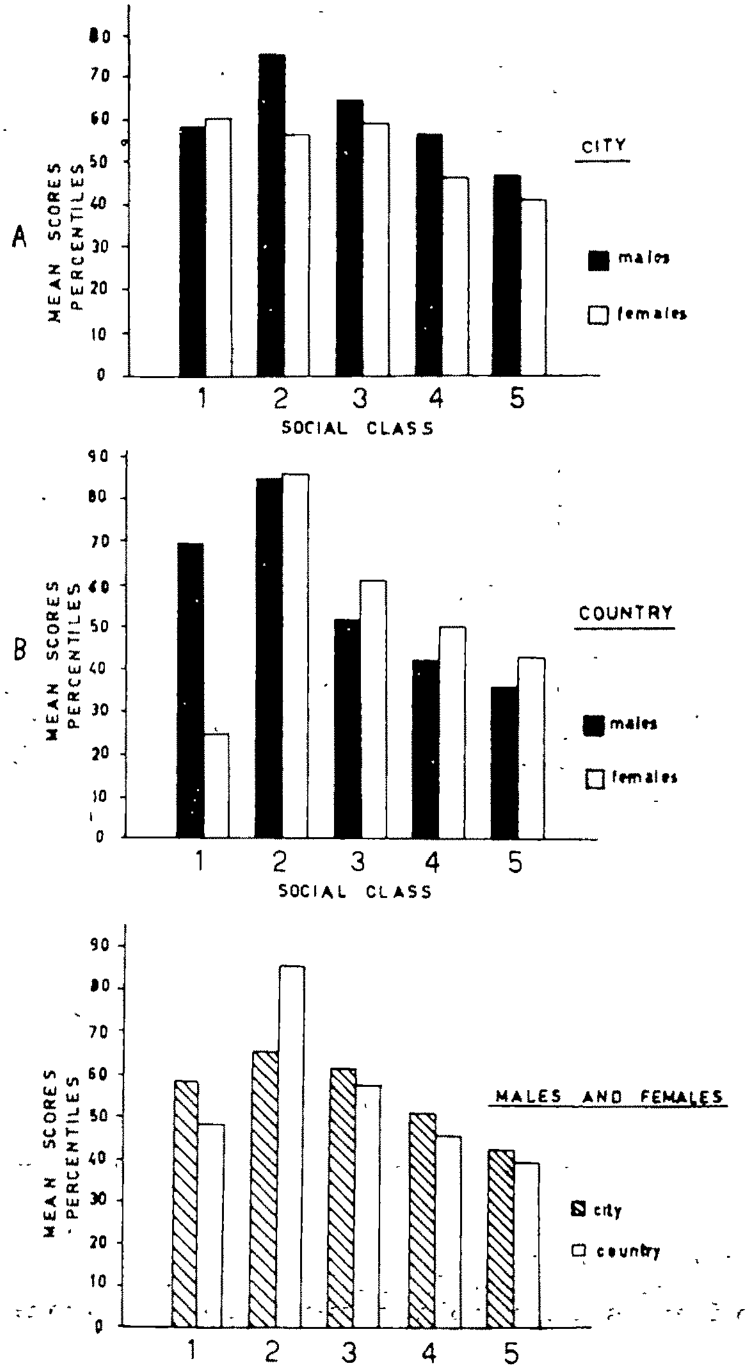


TABLE 3

COMPARISON BETWEEN MALES AND FEMALES WITHIN EACH SOCIO-ECONOMIC CLASS.

	Socio-economic class	Mean score		Differ'nce between the means	Degrees of freedom	t	p
		Males	Females				
City	1	57·40	59·86	2·46	27	·21	·80–·90
	2	74·29	55·50	18·79	31	2·07	·05–·02
	3	63·98	58·21	5·77	257	1·96	·05
	4	56·13	45·18	10·95	380	3·66	·001
	5	45·59	40·16	5·43	88	0·96	·34
Country	1	69·50	25·00	44·50	2	—	—
	2	85·00	86·00	1·00	3	—	—
	3	52·05	60·94	8·89	50	1·08	·28
	4	42·32	49·74	7·42	193	1·78	·08
	5	35·64	42·80	7·16	115	1·37	·17

There is again, the possibility that town boys, of whom a higher percentage than rural boys intend to proceed to classical junior high school (Table 4), are more motivated to perform well on the test. On the other hand, 38 per cent. of town girls, as compared with 22 per cent. in the country, also intend to proceed there and in this case the difference in plans, although less ambitious than for the boys, does not seem to have affected performance. Moreover, the mean results of those intending to proceed to classical junior high school are not significantly better than the means of the total populations for town and country. Table 5 shows the social class distribution of those who intend to proceed to classical junior high school. Although this table demonstrates that there may be need for greater equality of opportunity, there is nothing in it or in the other evidence presented to suggest that motivation of the boys was consistently affected by the situation.

TABLE 4

PERCENTAGE OF CHILDREN TAKING THE EXAMINATION FOR ADMISSION TO CLASSICAL JUNIOR HIGH SCHOOL.

	Males %	Females %	Total %
City	53·7	38·0	46·3
Country	22·5	22·0	22·3
Total	44·0	32·7	38·6

It might be that country boys devote less time to school work than town boys or town and country girls. Data given by the schools on absenteeism and time dedicated to homework each school day in each group indicate that there are no differences in number of days absent or time given to homework by boys.

TABLE 5

SOCIAL CLASS AND INTENTION TO ENTER THE CLASSICAL JUNIOR HIGH SCHOOL.*

Social Class		1	2	3	4	5
City	Males	93	88	78	42	24
	Females ..	100	81	64	20	1
Country ..	Males	100	100	65	20	16
	Females	50	50	69	16	2

There remained to consider the opinions and attitudes of the parents. A pilot study of forty families, ten from each group of boys and girls, town and country, shows no major differences in respect to the variables included in the questionnaire shown in Table 6.

TABLE 6

RESULTS OF QUESTIONNAIRE TO FORTY FAMILIES.

	Town		Country	
	Boys	Girls	Boys	Girls
1° In the fifth grade is it more important for boys or for girls to be able to read and compute correctly ?				
Boys	10	10	10	8
Girls	—	—	—	2
2° During the fifth grade, should boys have some kind of work to do, even if only in the home ?				
Yes	5	5	4	4
No	5	5	6	6
3° During the fifth grade, should girls have some kind of work to do, even if only in the home ?				
Yes	8	10	9	7
No	2	—	1	3
4° After the fifth grade, for whom is it more important to go on with schooling ?				
Boys	7	6	7	4
Girls	1	1	—	2
Boys and Girls equally	2	3	3	4
5° After the fifth grade, which school would you choose for your children ?				
Boys .				
High School	7	6	7	4
Business School	1	1	—	2
Trade School	2	3	3	4
Farming School	—	—	—	—
Post-Elementary	—	—	—	—
No School	—	—	—	—
Girls :				
High School	5	3	7	5
Business School	5	6	—	3
Domestic Economy School	—	1	3	2
Post-Elementary	—	—	—	—
No School	—	—	—	—
6° During the fifth grade, how much time a day on the average did your children spend on home-work ?				
About half-an-hour	1	2	2	—
About one hour	3	2	6	7
About one hour-and-a-half..	5	3	1	1
About two hours	1	3	1	2

---

*—Stated as percentage of children in each group.

Comparison of Tables 6 and 4 reveals that town children are much more likely to fulfil their parents' scholastic aspirations. Such frustration might lead to a type of defeatism that could conceivably have some effect upon the children, but, if so, it should be explained why boys should be more affected than girls.

We are left then, with the possibility that certain important aspects of social class, such as parental education and income, and the general condition of the household unit, which were not considered here,* might have been lower in the country than in town, causing a clustering of country children in the lower part within each social class. If this were so, it had no effect upon the females, but this factor might assume importance if one accepts the hypothesis that the male is more enviroment-vulnerable than the female (Tanner, 1961).

An informal enquiry among teachers suggests that in the country girls 'apply themselves better' to work in school. Further enquiry from country teachers elicited some opinions that in social classes 4 and 5 the boys frequently rose early to do physical work and were, therefore tired when they arrived at school.

The real nature of the differences in performance between town and country boys in social classes 4 and 5 can only be established by further studies. It seems important to undertake such work, as our findings may imply that, in this age of technology, the country boy might be less receptive than the girl to technical progress.

If the results of this study are confirmed, they should be taken into account in educational and social programmes in order that such a situation in the countryside may not, in the long run, become pathogenic.

ACKNOWLEDGEMENTS.—This study was made possible by support from the Grant Foundation, whom we wish to thank. We also thank the scholastic authorities for their wholehearted collaboration.

---

* Family size is an uncontrolled factor. We regret that we do not have data on this or on the places of the children in the family constellations.

IV.—REFERENCES.

ANASTASI, A. (1958). *Differential Psychology* (3rd ed.). New York: MacMillan.

ANASTASI, A. (1960). Cultural differences. *Encycl. Educ. Res.* (3rd ed.), pp. 350-358.

DAVIS, A. (1948). *Social Class Influences upon Learning.* Cambridge. Harvard Univ. Press.

HAVIGHURST, R. J., and JANKE, L. J. (1944). Relation between ability and social status in a mid-western community: I—Ten-year old children. *J. Educ. Psychol.*, 35, 357-368.

HAVIGHURST, R. J., and BREESE, F. H. (1947). Relation between ability and social status in a mid-western community. III—Primary mental abilities. *J. Educ. Psychol*, 38, 241-247.

RAVEN, J. C. (1956). *Guide to Using Progressive Matrices* (1938). Revised order 1956. London: H. K. Lewis and Co., Ltd.

TANNER, J M. (1961). *Growth at Adolescence.* First edition, Oxford: Blackwell Scientific Publications, 1955. Second edition, in manuscript, for publication, 1961.

Manuscript received, 6th October, 1961.

INTELLIGIBILITY OF SCHOOLS TELEVISION PROGRAMMES

By ALBERT ROYDS

(*Research Student, Education Department, University of Manchester*)

SUMMARY. Having viewed television lessons in elementary science, pupils did free-writing exercises in which they set down what they considered the lessons had been about.

Using word-counts as a measure of verbal intelligibility, it was found that the programmes were intelligible to the viewers, because the language standards of the lessons approximated to their own language standards.

Using teachers' estimates of the teaching importance and the teaching function of the contents of the lessons, it was found that the lessons were intelligible for subject matter also. This was indicated by the qualities of the statements recalled and set down, in the pupils' exercises.

I.—INTRODUCTION.

IN August, 1959, the first British series of television programmes for primary schools was introduced, by Associated Rediffusion, Ltd. The first nine science programmes of this series form the basis of the investigation under report. The aim of the investigation was to find the features of these schools transmissions, which appeared to influence their intelligibility to viewing pupils. Each of the programmes was of eighteen minutes' duration and was transmitted from all transmitters of the Independent Television Authority, except those sited in the North-West. The programmes dealt with the four ancient Aristotelian 'Elements': Air, Water, Fire, and Earth. (Two programmes each and a final recapitulatory programme to end the series.)

II.—THE LITERATURE.

Siepmann, C. A. (1958), writing in the United States said: "In Britain you are hard put to to find any written record with respect to television (in schools)." The same is just as true in 1962. Whilst there is a steadily growing body of information about the social effects of children's viewing of television in the home, academic study of schools television appears to be limited to the work of Vernon, M.D. (1953) and Roy, W. (1959). The former concluded that "it is difficult to achieve very close co-ordination of visual and verbal material and . . . this is essential to good comprehension and assimilation." The latter considers that, besides its subject matter, a successful schools programme requires a commentator with personality, using the right vocabulary and balanced programme techniques.

It is not surprising that, in the Report of the Nuffield Foundation Study, Himmelweit, Oppenheim and Vince (1958) suggested the need for teacher co-operation in further research into the potentialities of the medium.

Studies in the field of sound-only broadcasting have usually aimed at measuring the factors thought to be contributory to the comprehension of programmes. Schonell, F. J. (1935) thought vocabulary burden and the structure of the talks to be the main faults of programmes and that descriptive humanised material in a concrete setting was best understood. Vernon, P. E. (1950) after investigating thirty-five different, yet possibly influential, factors found close relation between interest ratings and intelligibility. Trenaman, J.

(1950) confirmed this relation of interest and understanding, when ' taking average scores,' but found variation at different educational levels Belson, W. (1952) commented on the influence on comprehensibility of: difficulty of words used ;' the talk's organisation ; and the listeners' background knowledge.

III.—PLAN OF THE INVESTIGATION.

The over-all plan of the present investigation was to compare the programme material, as it is recorded in the scripts-as-broadcast, with the material included by viewing pupils, in free-writing exercises written soon after their exposure to the programmes.

By the courtesy of Associated Rediffusion, Ltd., exact copies of the scripts-as-broadcast were made available in advance for classification and analysis. This was possible because their programmes to schools are transmitted from previously recorded video-tape.

Pupils' free-writing exercises were obtained from 383 subjects (193 boys and 190 girls) aged between 8 and 14 years, in fifteen schools distributed in all parts of England where the programmes could be received. Undue disparity was avoided for: urban and rural; socio-economic levels; age; and intelligence grades of the subjects.

All the subjects viewed all the nine programmes of the series, but each school group did free-writing exercises after viewing one specified programme only. Although each programme was transmitted at least twice, subjects were not exposed to more than one transmission of the programme about which they would write their exercises.

For the purposes of the investigation, it was felt that free-writing would be likely to reflect the more uninhibited reactions of the subjects. Vernon, P. E. (quoted by Trenaman, J., 1950) considered " freely written reproductions were preferable to objective questioning,'' Strang, R. (1942) arrived at a similar conclusion and Vernon, M. D. (1953) found that free-writing, with its associated mistakes, confusions and inconsistencies, gave the best evidence of degree of understanding.

(i) *Conduct of the Investigation.*

Previous to each transmission, a list of statements to be made ' from the screen ' was drawn up. This list was supplied to the head and class teachers of the schools where pupils would view and write about the programmes. Teachers were asked to evaluate each statement in the list, according to: (*a*) their estimate of its value in contributing to the avowed aim of the programme; and (*b*) their estimate of its teaching function. The estimates, each on a ten point scale, were totalled to give an evaluation, by professional teachers, for each statement in a programme.

Pupils viewed the programmes under the usual conditions for seating, room shading, etc., and care was taken to avoid any suggestion that the occasion was a special one. Thus, no indication was given, either before or during the transmission, that the written exercise would follow. There was no note-taking during the programme.

Following the programme, a fifteen-minute ' break ' was allowed, after which pupils re-assembled, under the usual classroom conditions and, having writing materials at hand, were told: " Suppose someone asked you what to-day's programme was about, write down what you would tell them.'' This instruction was not amplified or supplemented in any way. The free-writing was not limited, either for its length or for time taken in its production. When completed, the head teacher added information as to age and intelligence grading, to each exercise. Local authority gradings for its classification system were the criteria for intelligence.

(ii) *Treatment of the Data.*

Allen, W. (1952) estimated that an average listener only understands about a fifth of what he hears. Nias, A. H. W., and Kay, H. (1954) tested how far listeners' remembering depended on the type of material and on such variables as their occupation, age, sex, and intelligence. They concluded that story-type material was the most effective.

It seemed desirable, therefore, as a first step, to consider the degree of verbal intelligibility of the programme material. As other investigators in the sound-only field had found, the difficulty was to find suitable 'measuring rods.' Vernon, P. E. (1950) had adopted methods previously devised for the measurement of degrees of readability of printed material. In default of any obvious alternative, a similar method was adopted for this part of the present enquiry. Accordingly, word-counts of various parts of speech, etc., were made from (*a*) the scripts-as-broadcast, and (*b*) the pupils' free-writing exercises.

Before word counting could be undertaken, criteria had to be set up for each category. It was also necessary to examine the consistency of the words used in the scripts, as well as in the exercises. In the case of the scripts, consistency measures were found by correlating word counts for the upper and lower halves of the middle pages and then applying the Spearman-Brown formula. In the case of the exercises, the counts were made in the first and second halves of the exercise of the pupil in each batch whose I.Q. was 100 (or nearest thereto). The results are set down in Table 1 below :

TABLE 1

	Consistency Index	
	Scripts-as-broadcast	Pupils' Exercises
Active Verbs	0·852	0·673
Adjectives	0·796	0·725
Adverbs	0·841	0·743
Nouns (concrete)	0·882	0·769
(abstract)	0·653	0·314
(proper)	0·510	0·401
Numbers	0·812	0·521
Pronouns	0·641	0·680
Prepositions	0·561	0·327

In the word-counting, applied to pupils' exercises, only cogent writing was subjected to the count, i.e., the count followed the deletion of the erroneous matter.

Table 2 gives the data resulting from word-counts. Two other measures are also recorded, viz. : the ratio of syllables per hundred words (The 'Flesch Count') ; and the ratio of Words per Sentence.

TABLE 2

	Scripts-as-broadcast		Pupils' Exercices		Correlation
	Word Counts				
	Means %	S.D.	Means %	S.D.	r
Active Verbs	7·96	1·07	7·99	1·57	·44
Adjectives	6·21	1·08	4·62	1·22	·41
Adverbs	5·19	1·24	4·74	1·06	·29
Nouns (concrete)	8·05	2·59	4·77	2·68	·37
(abstract)	5·10	1·43	3·66	0·91	·03
(proper)........	0·75	0·46	0·51	0·08	·25
Numbers	0·28	0·47	1·20	1·16	·17
Pronouns	5·26	0·96	3·82	1·86	·35
Prepositions	4·93	1·27	4·71	1·20	·13
	Other Ratios				
Syllables per 100 words	123·21	4·81	130·83	14·2	·47
Words per Sentence....	11·74	1·73	12·40	1·81	·55

In the case of the Scripts-as-broadcast, the rate of delivery of the words spoken was calculated, whether they were spoken in the studio or were from film or tape. The mean result was 79·30 words per minute with S.D. of 12·80.

Consideration was next given to the factual content of the programmes and its degree of recall by the subjects.

Teachers had already assessed the ' from the screen ' statements according to their (*a*) teaching importance and (*b*) teaching function. Of these statements, some were recalled with varying degrees of accuracy and some were ignored, in the pupils' written exercises. There was, of course, some adventitious material found in the exercises. Measurements, given in Table 3 below, were

TABLE 3

	(i) Percentage of all scripted statements available for recall	(ii) Percentage of all scripted statements actually recalled
(*a*) *Teaching Importance.*		
Considered ' Major ' statement (Teacher-scaling : 41+)	36·4	7·19
Considered ' Minor ' statement (Teacher-scaling : 11—41) ..	58·4	5·80
Considered ' Incidental ' (Teacher-scaling : 0–10)	5·4	0·14
(*b*) *Teaching Function.*		
To aid reasoning	78·4	6·68
For dogmatic acceptance	31·1	3·20
Confirming previous knowledge ..	12·3	0·90
Englarging pupils' experience ..	54·0	3·94
Recapitulatory (same programme)	19·5	3·05
(*c*) *Type of Visual Support*		
Reality	33·1	2·85
Pictorial (still or film)	30·3	2·36
Diagrammatic (still or animated)	11·2	1·19
Model	2·1	0·17
Unclassified	0·19	0·04

made after erroneous statements in the pupils' exercises had been deleted. Table 3 (*c*) classifies the statements made in the programmes and their recall by pupils according to the type of visual 'support' which they were given in the presentation of the programme.

The relationship of the nature of the material recalled by pupils and their age groupings and intelligence gradings is set out in Table 4 below.

TABLE 4

Type of Statement recalled by Pupils	Pupils' Age-Groups Under 10	10–11	11+	Pupils' I.Q. groups High	Medium	Low
	%	%	%	%	%	%
Accurate and derived from the screen..	18·1	51·4	8·2	42·7	23·7	11·3
Accurate but not derived from the screen	1·0	4·3	3·1	3·1	3·6	1·7
Partially true and showing screen influence	0·5	0·4	1·7	0·2	1·0	1·4
Trivia	1·8	0·6	1·6	0·8	1·6	1·6
Erroneous	0·9	0·9	1·2	0·3	1·0	1·7
Accurate and derived from a previous programme	1·4	1·0	1·9	2·7	1·3	0·3

IV.—DISCUSSION.

From the data of Table 1, it may be inferred that the language of the scripts of the programmes was consistent and also comprehensible to the subjects, the three lowest rated of the vocabulary correlations being in respect of: proper nouns, abstract nouns, and prepositions. This is not regarded as invalidating the argument that the scripted material matched the levels of the subjects' comprehension. Further confirmation of this is shown when the nouns from the scripts-as-broadcast are tested against the frequency ratios assigned to them in the Thorndike Junior Dictionary. Table 5, below, gives the result of applying this test to all nouns which were used in the programmes with a scientific connotation:

TABLE 5

SCIENTIFIC NOUNS FROM PROGRAMMES.

	Thorndike's Gradings, " per million " from Junior Readers. 101 or over	99—50	49—25	24—10	Fraction	Unclassified
Percentage of Nouns in script	45·5	13·0	19·5	18·1	0·9	2·9

The foregoing data suggests that, both on the grounds of its comprehensibility, and its expression the vocabulary was intelligible.

The rate-of-delivery of spoken words has a higher value than subjective impression suggested during transmissions. It would seem that there is scope for useful research to find the optimum rate of delivery, for pupils of varying ages and attainments.

The device of using Teachers' assessments, as measures of teaching importance and of teaching function, appears to be justified, on the grounds that their evaluations were the result of their professional skill and, as the individual assessments were pooled, the over-all result was likely to be a close approximation to accurate values.

As Table 3 (*a*) indicates, the pupils' recall rate was higher in the case of statements of major teaching importance (ratio approximately 1 in 5) than for statements of minor teaching importance (ratio approximately 1 in 10). In all calculations concerned, with pupils' recall, repetitions of the same statement, in either scripts or exercises were disregarded, i.e., any statement was counted only once If, from the major statements in each programme, those which express elementary scientific principles (the lesson-aims of the various programmes) are isolated, the number averages 2·4 per programme and, for these, pupils' recalls average 1·72 (ratio 1 in 1·4). This suggests that the lesson-aims were satisfactorily attained.

Nevertheless, it should be kept in mind that the function of television in schools is wider than the inculcation of factual information. Table 3 (*b*) shows the degree to which these additional functions entered into the designing of the programmes. In calculating the percentages in Column (i), the same statement was counted under more than one heading, if it served more than one function in a programme. The pupils' recall ratios were all very close to 1 in 12, except in the case of recapitulated statements, where the ratios rose to 1 in 6.

One of the main functions of television in schools is to open up new and enlarging vistas of information and experience. It will be seen that enlargement of pupils' experience accounts for more than half the statements made ' from the screen ' and this is adequately represented in the proportion of those statements which are recalled.

There is scope for further research into the teaching functions, mentioned in Table 3 (*b*) and into the supporting visual illustrations, as listed in Table 3 (*c*). In the present investigation of the nine scripts-as-broadcast, the impact of dogmatic material was seen to be not much less than that of material to aid reasoning, but this is not to say that similar results would flow from a different series of scripted programmes Again, the classification " Confirming what is already known by pupil," is a somewhat arbitrary one, its criteria being teachers' pooled opinions. More research on both these matters would be rewarding. The effect of the recapitulation of subject matter is interesting. On the present showing, recapitulation facilities recall, as might have been expected. In the course of the extraction of this information, however, it appeared that there is a point in programme structure where recapitulatisn becomes subject to the law of diminishing returns. This, too, suggests the need for further investigation. Incidentally, if any need exists to justify the comparatively low ratios in the last column of Table 3, it must be said that these data are entirely the result of pupils' *written* expression, whereas one result of television in schools ought to be to motivate pupils' activity in experiment and very simple research.

It will be noted that the categories in Table 3 (*c*) are fairly uniform, as are also those of Table 3 (*b*). This suggests that there is considerable inter-action between the teaching function of statements and the visual support which they receive in the programme. That having been granted, it seems likely that the categories have been too grossly drawn to produce the maximum evidence about programme structure, and this too could usefully form the basis of further enquiry.

Even with this limited amount of data, it would seem that the subject matter of the programmes was intelligible to the subjects who were exposed to it. Regarded as a ' pilot ' research, the data point to the need for further investigation, with more refined categories and techniques for measuring inter-variability.

The data of Table 4 only serve to confirm what might have been suspected at the outset, namely, that intelligibility is a function of age and of intelligence grading. With more subjects than 383 obtained from urban and rural schools and grouped for socio-economic levels as well as sex and intelligence groupings, it should be possible to conduct an analysis of variance for five factors, in extension of the present enquiry.

V.—CONCLUSIONS.

(1) The transmitted material of the television programmes under review was verbally intelligible to the subjects who were exposed to it under normal school viewing conditions.

(2) One reason for the intelligibility was that its language approximated to the standard of the subjects' own language, as used by them in free-writing exercises, after their exposure to the programmes.

(3) The intelligibility of the subject matter of the programmes appears to be commensurate with the teaching importance and teaching function, as scaled by the teachers of the subjects.

(4) Extension of the present investigation, with refinements in the measuring devices and with a larger population, would be useful.

ACKNOWLEDGEMENTS.—Grateful thanks are tended to Associated Rediffusion, Ltd., for permission to make use of scripts-as-broadcast for the programmes Similar thanks are tendered to Mr. E. Stafford and Dr. F. W. Warburton, Tutors of the Education Department, The University, Manchester, for valuable help and guidance.

VI.—REFERENCES.

ALLEN, W. (1952). Question of intelligibility. *B.B.C. Quarterly*, 7, 3, 147.

BELSON, W. A. (1956). Technique of studying the effects of a T.V. broadcast. *J. Appl. Stats.* 5.

HIMMELWEIT, OPPENHEIM and VINCE (1958). Television and the child. *Nuffield Foundation Study*. 60.

NIAS, A. W. H. and KAY, H. (1954). Immediate memory of a broadcast feature programme., *Brit. J. Educ. Psychol.*, 24, 154.

ROY, W. (1959). A year of school television. *Bull. Cambridge Inst of Ed.*, 2, 1, 6.

SCHONELL, F. J. (1935). Broadcasting in the senior school *Brit. J. Educ. Psychol.*, 5, 299-322.

SIEPMANN, C. A. (1958). T.V. and our school crisis. *Dodd, Mead, New York*, 23.

STRANG, R. (1942). Explorations of reading patterns.

TRENAMAN. J. (1950). Understanding radio talks. *Adult Education*, Dec., 176.

VERNON, M. D. (1953). Perception and understanding of instructional T.V. programmes. *Brit. J. Psychol.*, 144, 116.

VERNON, P. E. (1950). Investigation into the intelligibility of educational broadcasts *B.B.C. Audience Research*, Nov.

(*Manuscript received* 19*th October*, 1961.)

THE RELATIONSHIP BETWEEN READING ATTAINMENT AND THE INCONSISTENCY OF ENGLISH SPELLING AT THE INFANTS' SCHOOL STAGE

By JOHN A. DOWNING

(Institute of Education, University of London)

Summary. The current investigation into the effects of providing infant beginners with books printed in a phonemically regular and limited alphabet with consistent spelling has revived interest in the proposition that the unsystematic spelling of English is a major cause of reading failure. Critics of traditional orthography suggest that it makes the development of the skill of reading too complex a task for the young beginner. The excessive number of variations of visual pattern are a hindrance in " Look-and-Say " methods, and the inconsistent relations between letters and phonemes are a barrier to success with ' phonics.' Previous tests of methods utilising systematic spellings at the commencing stage with a later transfer to conventional English orthography have had significant successes, but these researches were conducted before the days of objective testing and without comparisons with control groups. The results of experiments on children's reading of more or less regularly spelt words suggest at the least that the role of English spelling's inconsistencies in reading failure should be investigated further. The study of the effectiveness of the Augmented Roman Alphabet for the early stages of learning to read promises to provide more conclusive evidence on this problem.

I.—Is UNSYSTEMATIC SPELLING A CAUSE OF READING FAILURE ?

The view that the unsystematic nature of English spelling may be an important cause of reading failure has been advanced a number of times in the past 120 years, and several attempts have been made by educationists to devise special methods to overcome the confusion which it is thought this inconsistency of English spelling may create in the mind of the young child making his first attempts at reading. At least four important experiments with such methods have been carried out and in every case, very significant improvements in reading attainments have been reported.

In spite of their success, none of these methods has been generally adopted, and it seems likely that this is due to the fact that many teachers have not recognised that the conventional spelling of their native language may be an important cause of serious reading difficulty, or perhaps it is because adults' memories of how they first learnt to read at around 5 years of age are naturally somewhat faded. However, wider interest in the possibility of easing the difficulties has been aroused by Pitman's (1961) recent paper on " Learning to Read : An Experiment."

He suggests that there are pitfalls of two types in the conventional spelling of English, namely the lack of consistency in the visual patterns of whole words, and the lack of systematic relationships between the individual characters of the conventional alphabet and the phonemes of the spoken language.

The lack of consistency in the visual patterns of words is considered to be serious because a large majority of infants' teachers at some stage use one or other of the ' global ' teaching methods, i.e., " Look-and-Say ", " Whole Word," or " Sentence " methods, the chief feature of which is their common aim of developing associations between the *meaning* of the word and its *specific visual pattern* in print. Unfortunately, the traditional spelling of English does not provide those consistent patterns which are necessary if the young child is to establish a clear visual image of the whole word or sentence. The various ways in which the following sentence may be written, provide examples. The six words it contains have produced thirty-three different visual patterns.

any sentence can be an example

any can an example

Any Be An Example

Example

ANY SENTENCE CAN BE AN EXAMPLE

any sentence can be an example
Any Sentence Can Be An Example

It may be thought that a good teacher will avoid presenting the beginner with so many different ways of writing the same word, but even in the beginning stage when children are perhaps making their own reading books, there will be books in the class-room library, and most of these books will contain such inconsistencies. If the visual patterns presented to the child are so irregular, it is difficult to see how he can be expected to form the necessary associations between particular patterns and particular meanings. The following table of the twelve commonest words of the young child's speech, derived from Burroughs' (1957) list, shows that they may be written in at least fifty different ways ; of which no less than thirty-nine are common occurances.

GO			go	go	Go	go
DO			do			
IN	In		in		In	in
DOG	Dog	Dog	dog	dog	Dog	dog
ONE	One					
BIG	Big	Big	big	big	Big	big
IT	It		it		It	it
UP	Up		up			up
I					I	
OUT	Out		out			
AND	And		and	and	&	
A			a	a		

It seems a plausible hypothesis that the presentation of such a large number of different visual patterns for the single meaning of words makes the task of learning to read too complex for many young beginners.

The second type of spelling irregularity which, it is thought, causes difficulty in learning to read, is the lack of systematic relationship between the individual characters of the conventional alphabet and the phonemes of the spoken language which they are said to represent. This criticism of the conventional spelling of English seems significant because another common method of teaching young children to read is the " Phonic " method. Although this method was, at one time, unpopular with teachers, the majority now use it at least as an adjunct to " Look-and-Say." A purely ' Global ' method is uneconomic since 70 per cent. of the English language becomes accessible when a child has learnt only 220 different syllables, and, what is more important, children seem inclined to make a phonic approach to language themselves, searching for rules to link their life-long experience of the spoken language with the printed word which they have to interpret. However, this search for rules seems likely to be frustrated by the inconsistencies of conventional English spelling. This absence of system in the relationship between the characters of the alphabet and the phonemes of English is of two kinds ; firstly a single symbol represents a variety of phonemes and secondly a single phoneme often has a large number of alternative spellings.

The more serious difficulty is produced by the fact that one character may have more than one meaning. For example, if a child discovers that the words ' bone ' and ' lone ' have common phonemes, and that the letters ' o-n-e ' represent this sound, he may be inclined to apply this rule to other words containing these letters, where, in fact, the pronunciation is different, e.g., ' done .' If he modifies his rule on finding that other words are pronounced like ' done,' e.g., ' one,' his rule-seeking behaviour is frustrated again when he encounters the word ' gone.' This example is only one of very many which have been quoted by writers criticizing conventional spelling for its false implications of auditory relationships.

The other aspect of this inconsistent relationship is that one sound may be written in many alternative ways. The effect of this, it is thought, is to give false implications of auditory dis-relationships. For example, the same phoneme occurs in ' many,' and ' penny,' but the spelling of the phoneme is different. The same applies in ' cow ' and ' bough,' and many such examples may be found in English. Pitman has listed twenty-two ways of representing the sound ' i ' as in ' ice-cream ' :

A	ais-e	aisle
	ai	Port Said
	ay	McKay
	aye	ayes
E	ei	seismic
	eigh	height
	ey	geyser
	eye	eye

I	i	child
	ia	phial
	i-e	file
	ict	indict
	ie	lie
	ig	sign
	igh	sigh
	is	island
O	oi	choir
U	ui-e	guide
	uy	buy
Y	y	by
	ye	bye
	y-e	style

Although it is thought that this false implication of auditory disrelationships is less serious than the false implication of relationships, nevertheless, the burden of so many alternative ways of expressing the same sound may be heavy for young children, with the result that they lose confidence in their ability to cope with the learning task.

II.—Empirical evidence ; experimental alphabets.

Most of the evidence supporting the view that spelling inconsistency is a cause of reading difficulty consists in reports on the greater efficacy of special teaching methods designed to avoid presenting such inconsistencies to children *in the early stages.* A number of experiments have been conducted in Britain and the United States, and the teaching method has been basically the same in each case. A perfectly systematic notation is used at the beginning of the infant's school period and is continued until the child has acquired fluency in reading, when a gradual transfer is made to reading and writing with the traditional alphabet and spelling.

The American researches are described in detail in two theses by Crockett (1950) and Tauber (1958). One early experimenter was Zalmon Richards, who claimed that in 1844 he had taught reading with a phonetic alphabet in one-quarter of the usual time and with greater accuracy, and, at about the same period, a number of schools in Massachusetts were experimenting with 'fonotypy,' a phonetic alphabet designed by Pitman and Ellis. This was used for the beginning stages of learning to read followed by a transition to conventional orthography, and the children were said to have been as a result, " in several of the essentials of good enunciation and reading, *years* in advance of most children who had been taught in the old way."

The first large experiment took place in ten schools at Waltham, Massachusetts, between the years 1852 and 1860. Fonotypy was again used for all children's first experiences with printed English, each school being supplied with copies of the " First Phonetic Reader," printed in the fonotypic alphabet, to be used in ' daily enunciatory drill.' After six to eight months, when the children could read fonotypy fluently, they were helped through the stage of transition to reading print in the traditional alphabet in a series of lessons arranged according to the variation of words from phonetic spelling, the pronunciation of the traditional spelling being indicated by an example printed in fonotypy.

The results of the Waltham experiment are given in a report of a Committee of the American Philosophical Society in 1899, which says of fonotypy :

> " We tested it thoroughly for six or seven years in the town of Waltham, Massachusetts, which then had about eight-hundred children in the public schools. The effect upon the school life of the town was very marked. The saving of time in teaching the children to read and to spell enabled us to introduce exercises for the eye and the hand, thus cultivating habits of observation, skill in drawing and writing, and geometrical ability. The fonetic print corrected the brogue of the Irish children and the Yankee dialect of the American in a surprising manner. An improvement in the moral and intellectual tone of the schools was also noticeable, arising certainly in part from giving the children interesting reading instead of such absurd falsehoods as that of saying that ' sea,' ' you,' ' pea,' spells ' cup.'
>
> " Fears were expressed that this method should injure the pupils' spelling. In order to test that question, I took pains to procure, several times, lists of words which had actually been used in Boston, Roxbury, and other places, with the percentage of failures on each list. Springing these lists, without warning, upon classes of the same grade in Waltham, we always found our percentage of errors very much smaller than in other towns, sometimes I think only one-third as large. We also questioned each pupil in our high-school as to the amount of time which he or she had devoted in his or her whole school life to fonotypy and fonography. Comparing these times with the percentage of errors in spelling, by the same scholars, we found that those who had read the most fonotypy made the fewest mistakes."

Increased speed of learning to read was again the chief criterion of the success of this type of method in the even larger scale research into Leigh's transition alphabet conducted in the schools of St. Louis, commencing in 1866 and continuing for at least twenty years. In an official circular, March (1893), quotes the report of the St. Louis school superintendent on the success of the method. Leigh had over seventy letters in his alphabet, but, although this was a larger number than strictly necessary, it adhered to the rule established in the Pitman and Ellis fonotypy that one symbol must stand for only one phoneme. The first reader of the standard series in use in St. Louis at that time was transliterated into Leigh's alphabet, and this was followed by a transition to the traditional alphabet in which the second and subsequent readers were printed.

The children in the St. Louis experimental schools took half a year to get through the first reader compared with the usual whole year, and by the end of the first year many of them had completed the second book in the traditional alphabet. All had read at least half of it, and the brighter pupils had completed a hundred pages of the third reader, whereas by conventional methods the second reader was not completed until the end of the ***second year*** at school.

Harris (the school superintendent) concluded that one-and-a-half to two years could be saved in learning to read by the use of a systematic phonetic alphabet for the early stages. He, too, reported that it led to additional educational advantages, claiming that the children " took more pleasure in reading books and newspapers at home " ; their spelling was more correct ; their speech was clearer ; and they were more intelligent in their general approach to other school subjects.

A further test of Leigh's method was carried out at Boston with similar results, and Gladstone (1879) states that Sir Charles Reed announced at the

1876 Philadelphia International Exhibition that the system had resulted in a saving of half the time usually needed for learning to read, two years' work having been cut to only one.

The attempt to isolate the remainder of the skill of reading from the special difficulties of English spelling apparently was not tried in Britain until the twentieth century, for the earliest report of the use in schools of a method based on this principle seems to be that of Jackson (1914). He writes that Miss McCallum, an infants' teacher in a school at Cowdenbeath, had used a special alphabet very similar to the International phonetic one, and with a perfect one-to-one relationship between character and phoneme for the children's first year at school. She began with well-known nursery rhymes, starting from the meaning of the whole passage, then analysing it into words and finally analysing the words into their sounds. She seems to have included a good deal of speech training in her method, making sounds and saying, "Watch my mouth. How many shapes do you see?"

An important advantage she saw in this method was that it allowed her to use words for reading and writing lessons ***purely on the basis of the children's own vocabulary.*** There was no artificial selection of 'phonic' words and no artificial rejection of words difficult to spell. At any rate, Jackson found that "The reading was much more fluent, natural and intelligent than is commonly found in an Infant School." The outstanding result was that at the end of one year with the phonetic alphabet, "the children can read any matter presented to them in phonetic symbols, and can write the same to dictation."

Fifteen British schools experimented with another special "simple phonetic orthography," between 1915 and 1924, and the results are published in a pamphlet of the Simplified Spelling Society (1924). Although all these schools used the 'nue spelling' and the special reading books which utilized it for the early stages, they seem to have made the transition to books in conventional spelling at different times. Some fixed an arbitrary date for the transfer and this varied between a few months and one year. Other schools allowed children to make the transition on the basis of their individual progress in reading. The separate reports from each school are given verbatim in this pamphlet and an analysis of these shows that there was general agreement on the method's success in speeding up the process of learning to read, teachers' and inspectors' estimates of the time saved ranging from half-a-year to two years. A further result was that children found no difficulty with the transition from the special learning alphabet to traditional orthography, which finding was not only common to these experiments with 'nue spelling' but was also a feature of the results of all the other previous experiments reviewed above.

III.—EMPIRICAL EVIDENCE: INVESTIGATIONS CONTRASTING PHONETIC WITH IRREGULAR CONVENTIONAL SPELLINGS.

A number of novel research approaches to the problem of the relationship between the unsystematic spelling of English and reading difficulty were employed by Lee (1960). He attempted to relate the relative phonetic regularity of the spelling of different languages with the methods of teaching reading used in the relevant countries and with their standards of reading attainment. It seemed possible that 'phonic' methods would be more popular in those countries where the spelling had a more regular relationship with the sound of the spoken language, and that the wholistic "Look-and-Say" approach would enjoy greater favour where the spelling was not so systematic. Lee showed that, although the terminology and methodology of teaching methods overlapped considerably from country to country, it was obvious that " a clear association between methods and ' regularity ' does not exist." For example, Finnish and Turkish are both among the most regularly spelt languages from the phonetic point of view, yet Finland uses the ' synthetic ' or ' phonic ' approach which Turkey has abandoned.

Lee found that comparisons between reading attainments of different countries with more and less regularly spelt languages were of doubtful validity, chiefly because there were no generally acceptable standard measures of achievement, but also because many other factors in the educational situation (e.g., age of commencement of formal schooling) vary and are thus uncontrolled.

The view that irregular spelling is an important cause of reading difficulty suggests the hypothesis that irregularly spelt words should be more difficult to read than regularly spelt ones. To investigate this, Lee developed an objective method of scoring the degree of irregularity in the spelling of English words, a regularly spelt word being defined as one in which every sound is spelt as that sound is most commonly spelt in the language. Thus, the irregularity score of a word rises as the spellings of its sounds move away from their most usual forms. In his two experiments, Lee classified the words which had to be read by the children into two categories: *more* regular and *less* regular, a dividing line being drawn arbitrarily at a point on his scale of spelling irregularity. He went on to analyse the errors made by 275 children in reading aloud selected passages of prose, and found that there was a slight tendency for the younger children (under 11 years) but not the older ones (over 11 years) to make more errors on the more irregularly spelt words than on the more regular ones, but that the difference was too small to warrant a definite conclusion.

In a further experiment on 172 junior school children (ages 7-11), the same author compared the errors made with twelve more regularly spelt words and twelve less regularly spelt words. Commenting on his results, he points out that five of the six most often correctly read were from the list of more regular words, while five of the six words producing the highest frequency of error were more irregular words, but Lee says, " As far as other words (in his list) are concerned, there is no correlation between the irregularity of their spelling and the extent to which they are successfully or unsuccessfully read," and later concludes from these results that irregular spellings are " by no means a major cause of reading difficulty among the children who took part in the enquiry."

Although Lee publishes no coefficient himself, the correlation between reading errors and spelling irregularity can be calculated from his data and turns out to be $+0{\cdot}37$ (Spearman's ρ), which is short of significance at the 5 per cent. level of confidence.

However, the validity of Lee's conclusion must be questioned, for he failed to control an important variable in his experiments. These children were between 6 and 12 years of age who had already learnt to read. Therefore, the effects of the irregularity of the spelling may have been obscured by the fact that his test words varied not only in spelling irregularity but also at the same time in frequency of occurrence in the children's reading and conversation experiences.

Table I sets out the relationships between Lee's results on spelling irregularity and reading accuracy on the one hand and Burroughs' (1957) finding on the frequency of these words in young children's speech. (Where Burroughs has not kept a separate record for a particular word figures from the report of the Scottish Council for Research in Education (1948) have been used.)

TABLE 1

SPELLING, IRREGULARITY, FAMILIARITY, AND READING ACCURACY.

Lee's Category of Regularity	Lee's Order of Accuracy in Reading		Burroughs' Percentage Frequency of Occurrence	
	No. of words in top half (i.e., more accurate)	No. of words in bottom half (i.e., more inaccurate)	Words used by less than 6% of children	Words used by more than 6% of children
A (More Regularly Spelt Words) ..	7	5	10	2
B (More Irregularly Spelt Words) ..	5	7	4	8

It is clear from this table that Lee's ' more irregular ' category was made up of words which are more frequently used by young children than the words he selected for his ' more regular ' category, and it seems clear that the correlation between reading errors and irregularity of spelling would have been greater had his test not been so weighted against the hypothesis that unsystematic spelling causes reading difficulty.

In discussing Lee's data it must also be noted that he did not test for individual differences or for differential effects due to the fact that different age groups were involved in the total sample tested. Allowance also needs to be made for the fact that the experimental design included the presentation of the twenty-four test words always in the same order—instead of being randomized or balanced.

In a third experiment, Lee dictated ten *nonsense* words to 256 children between the ages of 6 and 12 years, to discover whether children had formed associations between phonemes and spelling. He found that there was not a great deal of correspondence " between the ways in which children had written the sounds in these nonsense words and the ways in which they are usually written in English," and concluded that " the children had formed no mental association between speech sounds . . . and the most frequent ways . . . of representing them." This is not a surprising result, but it is difficult to follow the author's subsequent reasoning that because this association has not been established there is no evidence for such a " tendency or demand in a young reader's mental make-up," for one could hardly expect the development of such associations in young children if the actual relationships between spelling and phonemes are so objectively obscure as to provide no opportunity for the necessary learning to occur.

IV.—CONCLUSIONS FROM PREVIOUS RESEARCH.

Lee's pioneer contribution is important for he has marked out a number of interesting avenues of exploration in this problem. Although his failure to control the familiarity of the words for the children makes it impossible to agree with his view that " the lack of a regular correspondence between sounds and letters has been shown to be a minor cause of reading difficulty, particularly where isolated words were concerned," his own admission that " complete beginners and the very worst readers were not and could not be included in this investigation " stands as a challenge to other researchers.

Neither do the other studies reviewed above on the effects of introducing a systematic alphabet with consistent spellings for infants in their first months at school provide conclusive evidence that the inconsistencies of English spelling are a major cause of reading failure, for the experiments were insufficiently controlled and lacked objective testing. The reports consist chiefly in subjective anecdotes of the children's superior reading attainments during and after their use of the regularized spellings, without any real comparison with a parallel control group of children learning with conventional spellings. Nevertheless, anecdotal, subjective and uncontrolled as they were, reports of such great improvements in reading attainment seem to indicate the urgent practical need for these experiments to be repeated under conditions of more rigorous scientific testing.

V.—A NEW ALPHABET INVESTIGATION : THE " AUGMENTED ROMAN ALPHABET " IN BRITISH SCHOOLS IN 1961.

Concern over the allegedly harmful effects of traditional English spelling on children learning to read continues to give impetus to the experimental modifications of the form in which the printed language is first presented to young children. On both sides of the Atlantic small studies are currently under way. Dr. Helen Bonnema is using ' dikshenerikee,' a regular phonetic alphabet at a school in Denver, Colorado, and in the South of England children at two infant schools are learning to read by Jones' " Colour-Story Method "—a logical system of forty colour-symbols, each representing a distinct sound which the children learn incidentally through stories which provide concrete images of the abstract phonetic concepts.

An interesting development in this field is the study of the use of the " Augmented Roman Alphabet " in the early stages of learning to read of children in British infant schools, which is being conducted by the present writer. In 1961 and 1962 a number of schools is expected to introduce this phonemically regular alphabet and its consistent spelling system for their intake of five-year-old children. The whole of the most commonly used infant reading scheme—basic books, supplementaries, and apparatus—has been printed in Augmented Roman (A.R.), as well as a number of attractive story books for the class-room ' book corners,' home use, and for local public libraries.

Objective comparisons between the A.R. experimental schools and matched control schools may provide more conclusive evidence as to the role of spelling in children's difficulties in learning to read, as well as testing the effectiveness of what might be called a ' two stage ' method of helping children to learn to read—a ' two stage ' method, because at first, the skill of reading is separated from the peculiar problems of the spelling of English, and only after the child has completed the course of A.R. books and developed the skill of extracting meaning from print is he presented at the second stage with the special difficulties of English traditional orthography.

It is important to understand that the major aim of the A.R. research project is to test the effectiveness of this ' two stage ' method in raising the level of young children's reading attainments along with increased confidence and enthusiasm. Failure to appreciate this purpose of the investigation seems to explain how Wijk (1960), in a rather strange document, came to make his premature criticism of the research before it had begun !

Wijk (1959) had earlier proposed a *reform of English*, "Regularised English," which was inappropriate for the present investigation since it retains a great number of the inconsistencies of the traditional spelling of English which have to be removed if the skill of reading is to be divorced from these peculiar difficulties at the beginning stage. Wijk's original approach was linguistic rather than pedagogical, and his system might be useful to students of language as such, by providing a regular system of exceptions which might be said to justify some of the inconsistencies.

The Augmented Roman Alphabet, an adaptation of the original designs of Sir Isaac Pitman and A. R. Ellis, in contrast, is specifically designed for the purpose of helping children to learn to read in the beginning stages and, therefore, priority has been given to fitting the learning task to the child's ability at this early stage of development. The table below shows its forty-three characters and their phonemic values. It has been designed to remove the

THE AUGMENTED ROMAN ALPHABET

Number	Character	Name	Example	
1	æ	ae	ræt	(rate)
2	b	bee	big	(big)
3	c	kee	cat	(cat)
4	d	dee	dog	(dog)
5	ee	ee	seet	(seat)
6	f	ef	fill	(fill)
7	g	gae	gun	(gun)
8	h	hae	hat	(hat)
9	ie	ie	ieland	(island)
10	j	jae	jieant	(giant)
11	k	kae	kit	(kit)
12	l	el	lamp	(lamp)
13	m	em	man	(man)
14	n	en	net	(net)
15	œ	oe	bœt	(boat)
16	p	pee	pig	(pig)
17	r	rae	run	(run)
18	s	ess	sad	(sad)
19	t	tee	tap	(tap)
20	ue	ue	fue	(few)
21	v	vee	van	(van)
22	w	wae	will	(will)
23	y	i-ae	yellœ	(yellow)
24	z	zed or zee	fizz	(fizz)
25	ƨ	zess	houƨeƨ	(houses)
26	wh	whae	when	(when)
27	ch	chae	chick	(chick)
28	th	ith	thaut	(thought)
29	th	thee	the	(the)
30	sh	ish	shepherd	(shepherd)
31	ʒ	zhee	juʒ	(judge)
32	ŋ	ing	siŋ	(sing)
33	ɑ	ah	fɑr	(far)
34	au	au	aut	(ought)
35	a	at	appl	(apple)
36	e	et	egg	(egg)
37	i	it	dip	(dip)
38	o	ot	hot	(hot)
39	u	ut	ugly	(ugly)
40	ω	ωt	bωk	(book)
41	ω	oo	mωn	(moon)
42	ou	ow	bou	(bough)
43	oi	oi	toi	(toy)

difficulties of traditional orthography in two respects. First, the variations in the visual patterns of words are removed and secondly, the inconsistencies in the auditory relationships between the characters and the phonemes they represent are abolished.

To achieve consistent visual patterns for whole words, each letter has only one form. Lower-case characters only are used and capitalization is achieved by using a larger form of the same lower-case shape, lower-case being chosen because it is believed that the ascenders and descenders provide a maximum of discrimination in the whole word pattern and that they are used more than upper-case letters in other books which children may see. It is hypothesized that this simplification should make it easier for the beginner to establish clear visual images of the patterns of whole words which he has to associate with particular meanings.

Both kinds of difficulty with the auditory relationships between characters and phonemes that are found in the traditional spelling of English are attacked by the designers of the Augmented Roman Alphabet. In A.R., each symbol[1] stands for one phoneme only, and it is alleged that this will allow the child to discover rules linking the printed letter and his past experience of the spoken language. The other difficulty of traditional spelling has not been entirely removed by A.R., for it still permits alternative ways of representing the same sound, although the number of variations is very much reduced from approximately 2,500 to 88 ways of representing the forty plus phonemes of English (counting the doubling of characters, e.g., *skill*, *skilful*, and *skilfully*, as alternative characterizations of the phoneme).

Pitman has departed from what might be considered the ideal of a perfect one-to-one relationship between letters and phonemes, because he considers that it is necessary to make a compromise between more complete simplification and the need for the child to make an easy transition from A.R. to the traditionally spelt English words which he must meet later. It is for this latter purpose, that more characters than are strictly necessary have been retained in A.R. For example, 'c' and 'k' represent the same phoneme and so do 'z' and 'ƨ,' and some English speakers would regard 'w' and 'wh' also as representing but one phoneme. For the same reason, double letters are retained, e.g., in 'letter,' 'back.' The design of the Augmented Roman Alphabet has been influenced by the need to ease the transition to traditional orthography in other ways. The above table shows that there are nineteen augmentations, but that most of these bear a resemblance to the characterization of the traditional alphabet. For example, fourteen of them are digraphic and the digraphs that have been chosen are those which Ripman and Archer (1948) showed to be the most common in the representation of the speech syllables of English, e.g., 'ch,' a most common digraph for a phoneme which lacks its own character in the traditional alphabet, is represented by 'ch' in A.R.

The problem of transition again influenced the design of the form of the characters in that, as far as possible, their 'top coast-line' has been left undisturbed. It was thought important to take this into account because of the need of the fluent reader to be able to continue to skip rapidly along the lines of print, paying attention only to the 'top coast-line' and the length of the words. If this has been achieved, Pitman's argument that transition may be regarded as more of an adjustment than a re-learning should be justified. In any case, as has been noted above, in all previous experiments of this type, no difficulty was found at the transition stage and it seems likely that this was due to the other incidental experiences which children had of the traditional spelling outside the class-room. In addition, by the time the child is fluent in reading, his past experience of the content of written English sentences enables him to fill in any ' gaps ' on the basis of context.

---

[1] The term "symbol" embraces combinations of letters, e.g. the digraph "er".

Summing up, Pitman hypothesizes that the Augmented Roman Alphabet helps the " Look-and-Say " approach because it makes the visual patterns invariable and thus produces stable visual images, and it helps the phonic approach, because each symbol stands for one single phoneme only and the number of alternative spellings for the forty plus phonemes of the English language has been considerably reduced, thus allowing children to discover rules linking the letters of the alphabet and the sounds of the spoken language. Furthermore, the design of the alphabet is such that any difficulties in the transition should be reduced to a minimum.

To return to the problem which this article is concerned with, the investigation of the Augmented Roman Alphabet should provide critical evidence on the relationship between English spelling and reading difficulty at two points. First of all, the most marked effects should be apparent while children are still using only books printed in A.R., and secondly, significant data should become available during the process of transition from A.R. to books in the traditional spelling. Only on the basis of such objective evidence can we be sure of reaching a correct solution of the problem stated by Inglis (1948) in his review of work on the early stages of reading: " The teacher of reading to English-speaking children must continue to face difficulties arising from the inconsistencies of letter and sound and from the vagaries of spelling."

VI.—REFERENCES.

BURROUGHS, G. E. R. (1957). *A Study of the Vocabulary of Young Children.* Edinburgh : Oliver and Boyd.

CROCKETT, W. D. (1950). *Towards More Rational Spelling.* M.Ed. Thesis, Tufts College.

GLADSTONE, J. H. (1879). *Spelling Reform from an Educational Point of View.* London : Macmillan.

INGLIS, W. B. (1948). The early stages of reading : a review of recent investigations. In *Studies in Reading,* Eds. Scottish Council for Research in Education. Edinburgh : Neill.

JACKSON, R. (1914). Phonetics and phonetic tests in the teaching of reading. *Miscellanea Phonetica,* 1, 33-37.

LEE, W. R. (1960). *Spelling Irregularity and Reading Difficulty in English.* London : National Foundation for Educational Research in England and Wales.

MARCH, F. A (1893). *The Spelling Reform.* U.S. Bureau of Education, Circular of Information No. 8.

PITMAN, I. J. (1961). Learning to read : an experiment. *Journal of Royal Society of Arts,* 109, 149-180.

RIPMAN, W., and ARCHER, W. (1948). *New Spelling.* London : Pitman.

SIMPLIFIED SPELLING SOCIETY (1924). *The Best Method of Teaching Children to Read and Write.* London : Pitman.

TAUBER, A. (1958). *Spelling Reform in the United States.* Ph.D. Thesis, Columbia University.

WIJK, A. (1959). *Regularised English.* Stockholm : Almqvist and Wiksell.

WIJK, A. (1960). *' gae, foot, dee ' spells* (' *good* '), *' yae, brood ' spells* (' you '). Stockholm : Seelig.

(*Manuscript received, June 16th,* 1961.)

SYMPOSIUM—EDUCATIONAL RESEARCH TO-DAY

I.—PRIORITIES IN EDUCATIONAL RESEARCH.

By E. A. PEEL.
(*University of Birmingham*)

AT a time when all concerned with education are aware of the need for investigation into many educational problems and new developments, I think we should look into policy and planning as regards present day research.

In my opinion, policy can be stated very simply. We need *more*, we need a *change of emphasis* and we need the *material and financial means* by which to do these things.

Education today poses a mounting number of unanswered questions which need objective study. Thus, the extension of facilities for education, particularly at the further and secondary stages, has sent a wave of problems back along the line to the primary stage, as, for example, in the learning of mathematics. The demands of modern industry and commerce, the increased load thrust upon the individual by a more complex world and an incessant drumming of communication, all ask more of a young person whose attainments and judgments may seem relatively to be found wanting.

Now all these problems involve us with questions as to the *progress* of our pupils and students—with educational processes in a wide sense, including learning and maturing and the bearing on them of teaching and organisation and of out-of-school influences.

As to the need for a change of emphasis, let me give two instances. Since the war we have spent a lot of effort on increasing the efficiency of secondary selection. This was historically and psychologically appropriate. But the allocation of pupils does not, in my opinion, constitute the *major* problem any more. This is not due solely to improved methods of selection, but also to the more flexible and comprehensive organisation of secondary education. There are still problems in this field centring around social factors and motivation, but I think we research workers should now turn much more to study *processes*. Education does not end when we have made a reasonable choice of the children. It is only the beginning. The important questions today concern the effectiveness of education. We can now investigate these problems with more confidence since we have better means for assessing understanding and learning.

Next, there is the aimless survey type of inquiry, whether by questions or by tests. I am not against surveys and longitudinal studies which are carried out to test clearly stated hypotheses, and in which there is a real intention to do something about it when all the information has been completed and computed. But the spirit of aimless surveys is often merely ' Let's see ! ' and that is where it ends. It is curiosity of a low order. A survey, like a test, is a servant, not a master.

The planning of research on learning and teaching problems compels us to ask the questions :

How best to orientate the research ?
How best to break up the topics ?
Where it can best be done ?

I use two yardsticks to guide me in answering the first question. First, whether the purpose is to investigate current practice or to look forward. Thus, we may wish to inquire into the effects of streaming in primary schools, or the effects of organising secondary education comprehensively upon the standards of its pupils. On the other hand, we think forward when we wish to try out new ideas, which may have arisen from research results or experienced observation, such as programmed learning.

Secondly, I like to settle whether the research is to be backroom (fundamental), concerned with a deeper study of processes, *or* whether it will be carried out in the school working setting. Some of the most helpful suggestions to classroom practice have come from backroom inquiry. I use ' backroom ' in a figurative sense in preference to the more pretentious term ' fundamental ' and avoid altogether the misleading adjective ' pure.'

Let us see how such backroom work, which is also forward looking, may have quite unexpected value in the working situation.

I wished to investigate the positive motivation of school learning (quite apart from external incentives like rewards and punishments). I fixed on curiosity and interest as being important carriers of learning. Of many questions asked in the research one was, ' *How is curiosity maintained and even increased* ?' we used strange objects and pictures—certainly not classroom events—to measure curiosity. We confirmed that progress depends on novelty pacing complexity. It was at this stage that I realised how important this could be for the compiling of exercises in text books and for the rate of the steps of a lesson. I looked at some text books on science and mathematics and found them wanting on account of uneven progression of exercises and allied steps in the exposition. I mentioned this ' pacing ' idea to American colleagues, experts on programmed learning, who told me that this is something they try to control very carefully in their productions. Whatever happens the pupil must not fail and leave the field. But the matter did not end here. We can overdo novelty. I had a research student working at the time on the effect of unfamiliarity of content etc., on problem solving in school mathematics. Here is an instance :

(1) If £1 will buy 96 packets of biscuits, how many packets should I receive for £2 5s. 0d. ?
(2) If £1 is worth 96 Egyptian piastres, how many Egyptian piastres should I receive for £2 5s. 0d. ?

The second gives much more trouble than the first.

This is a research carried out in the school setting. It finds common ground and is ' explained ' by the first, namely that here we have novelty *outstripping* complexity. Neither novelty or complexity must dominate. Novelty must just be a ' pacer.'

We require research on *processes* in the pupil and the effect schooling has upon them. Now, how to break up this objective ?

I would put high in the priorities studies of school learning, concept formation, thinking and reasoning, motivation, and the growth of values and morale. And all these should be considered in relation to teaching, curriculum and school organisation.

Here are some topics to illustrate the possibilities. We may study verbal learning in the backroom and language learning in the school setting. The direct method of teaching foreign languages and the teaching of English as a foreign language offer many leads. The results of research into verbal learning in a laboratory setting confirm the superiority of many of the direct practices of up to date language teaching methods.

Now that we have better means for assessing levels of thought and suggestive theoretical bases, we can break through into the adolescent period of thinking which has scarcely received any serious attention at all in the school setting. Backroom research gives us a picture of the growth of thought. Work on the school side in relation to physical science and mathematics can be extended to other school subjects. We have research students investigating thinking and the growth of concepts in history, geography, biology and general comprehension and criticism. There may be a general tendency in the early secondary school years to assume thought is more advanced than it really is, and one suspects sometimes in the upper parts of some schools often enough is not done to capitalise on the maturity of thinking that is possible, although one often gets the opposite tendency as well. Once we have a picture of the growth of thought as revealed in the study of school subjects, we can then improve the teaching and learning. Learning should pace thinking—but it must not get too far out in front. I have used the pacing analogy twice—first for improving motivation and now for bringing about understanding and achievement. Education is pacing—the pacing of the young by the mature.

Values and morale form the intangibles of education. But education fails completely if it has not brought about some real growth of standards. These can be studied both in the backroom and the classroom. The inquiry by Miller, published recently, on *Values in the Comprehensive School** showed how useful this kind of research might be. Comprehensive education was shown to achieve something not so marked in tripartite secondary education. The main findings are a promising augury for the idea of comprehensive schooling. The morale of the less able pupils was raised very much without loss in that of the abler pupils. There was, indeed, evidence that the abler pupils sometimes gained in attitudes and interests.

Teaching, like curriculum and organisation, is on the side of the influences bearing on the growing pupil, but it is so specific and so closely connected to the pupil's learning that we must look at it seriously as requiring research on its own account. Teaching that brings about learning with understanding and makes generalisation possible, guided problem solving, the use of mass media and programmed learning—all need objective scrutiny.

In the backroom we can profitably study verbal learning, reinforcement, spacing of drill, and direction of thinking by laboratory means. Teaching can gain from these detailed studies of the pupil's mental processes.

Now I turn to my last question, Where can these things best be done? I want to comment on the roles of the University Departments and Institutes of Education, the National Foundation for Educational Research and the Ministry, but I do not forget the place of local authorities, the teachers' organisations and the training colleges.

The University Departments and Institutes have the greatest responsibility. They could best take on the backroom, forward thinking research, as well as classroom projects. They can and should attempt to break new ground, particularly where the returns might be 'chancy.' They don't *have* to produce immediate results. They should be the source of trained personnel and those with expertise and some vision.

---

* T. W. G. MILLER (1961). Values in the comprehensive school. *Educ. Monog. V.* Birm. Univ. Inst. Educ.

But in some respects they are not all equipped to do this. Not all of them have sufficient earmarked funds. Institutes should concentrate more upon developing central research facilities and Departments might well look beyond initial training. Research staff making up a team of specialists, which should include top and senior personnel, and strong higher degrees and advanced diplomas sections all make for progress in educational research. Some of the most interesting work at Birmingham, Manchester and London has been in the form of co-ordinated research carried out by advanced degree and diploma students. University Departments of Education and Institutes of Education must look a lot further into research 'possibilities and the training of research personnel. The Foundation must, under its present system of finance, concern itself mainly with safer projects, elaborating and testing out the more promising of the probes put out by the University Departments. Its national relations with the Ministry, with local education authorities, as producer-consumers and with teachers, lend themselves to such work. But I am not saying it should not have the 'small backroom.' If I were running it I should insist on this function.

One of the Foundation's important functions should be communication. The weak link in all that I have said is in getting the findings to the teachers, but the National Foundation for Educational Research, through its publications, is making good progress in this respect.

This general division of labour puts the University centres and the Foundation in useful complementary roles.

All this needs money—more than we now get through U.G.C., and the contributions to the Foundation. We can make appeals, often successful, to the larger foundations and trusts, but why can't we have something like a D.S.I.R. or an M.R.C. with financial power at its elbow? Why not a D.E.R. or an E.R.C. with power to encourage and promote research? Such a body could help in advice, co-ordination (particularly of large projects like, say, the teaching of Mathematics) dispensing of funds, information, but not *direction*. Research ideas have an essentially individual element.

The essentials of this New Order are as follows:

More research of good quality.

A change of emphasis to take account of the processes of learning and growing up.

More adequate material and financial backing.

II.—CO-OPERATIVE RESEARCH IN INSTITUTES OF EDUCATION

By G. E. R. BURROUGHS
(*University of Birmingham*)

SUMMARY. This article speaks of the particular responsibility which Institutes of Education have of carrying out research on a collective basis in co-operation with lectures and teachers outside the universities. It then goes on to describe the attempts to do this in one particular Institute of Education, pointing out the difficulties met as well as speaking of its successes.

I.—INTRODUCTION.

WHEN the establishment of Institutes of Education began in 1947, particular thought was given to the manner in which they could best provide services to lecturers, teachers and other educationists in their Institute areas. All of them, in addition to their responsibilities for the co-ordination of teacher training, provided libraries, various full-time and part-time courses and facilities for the exchange of educational information at different levels. The purpose of these developments was to strengthen the links between educationists, particularly those in training colleges and schools, on the one hand and the universities on the other. In considering this provision of services the Institute of Education of the University of Birmingham was of the opinion that the contribution which the universities were particularly fitted to make was in the field of research and it has consequently always placed a special emphasis upon it.

Educational research has generally been the pre-occupation of individuals in or associated with the universitits. There are distinguished exceptions to this statement in the activities of such bodies as the Scottish Council for Research in Education and, more recently, the National Foundation for Educational Research and on a smaller scale, of groups like the Surrey Teachers' Research Group or *ad hoc* bodies set up periodically. This does not, however, lessen the fact that generally speaking few facilities have existed for those educationists who, by inclination or through limited availability of time, wished to carry out research collectively. As well, therefore, as providing funds and facilities for individual research at a high level, this Institute decided to provide further facilities for and to experiment with collective research conducted outside the university. It is of this Institute's experience in this field that this article principally deals. In deciding to develop a research organisation., the Institute was fortunate in having a firm base from which to operate. The University already had, through its Education Department, a reputation in the field of educational research ; on the foundation of the Institute, a strong Department of Research was established as an integral part of it which could take advantage of this existing experience. Indeed, the Head of the Department of Education was asked to become Head of the Institute's Research Department and this close relationship has always existed, an arrangement which enables academic staff to be easily deployed at one time in the Institute and at another in the Department. Thus, the writer, who was appointed particularly to develop co-operative research in schools and colleges, was appointed to the Department of Education to work within the Research Department of the Institute of Education, an arrangement which works admirably even though it is difficult to explain to visitors from abroad. It will be appropriate to stress at various points in this article what appear to be salient features in the growth of this special work as well as mistakes which have been made. These, it is hoped, may be of help to others We may say immediately that two important points were initially recognised; the desirability of rooting the new work in previous research experience and in the closest contact with the University, and the need to appoint someone with primary responsibility for it. Since previous research activity and the atmosphere which a University research team creates is not immediately available to all, the first of these must be regarded as a desirable condition and many would think it essential; the appointment of an appropriate full-time research adviser and organiser is certainly an essential one.

II.—INITIAL RESEARCH ORGANISATION.

The work of providing facilities for collective research for those educationists who had, perhaps, little research experience and certainly too little time to carry it out individually, was begun in 1950. It was decided that the initial major effort would be in the training colleges rather than the schools. Within the Institute there already existed groups of training college lecturers who met periodically to discuss matters of common interest. These Boards of Studies were identified with and by the academic subject with which their lecturer-members were principally concerned. Similar bodies exist in almost all Institutes of Education. The first attempts to introduce research topics were made through these Boards, these attempts following a period of campaigning in the training colleges themselves. During the next five years some good and interesting pieces of work were completed mostly of a survey nature. The Board of Studies for Modern Languages, for example, carried out and published a valuable survey of modern language teaching in secondary modern schools at a time when information under this heading was almost totally lacking.* The Board of Studies for Mathematics, too, made some interesting analyses of errors commonly made by children and by students in various branches of their subject. The Board of Studies for Housecraft and for Education, the latter always a very active group, also carried out surveys. The Education group in particular, adapted the child study techniques in which their students were being trained so that these might contribue to the research effort. Nevertheless, it became apparent that this method of working through Boards of Studies was not wholly successful. This was due essentially to the fact that the Boards of Studies by their constitution and terms of reference, were not primarily groups created for or interested in research. Thus, by the time they had dealt with their regular business of considering syllabuses, methods of examination or listening to and discussing some paper on their subject, tea-time was approaching. Against this, research often seemed a poor alternative. It was only a minority of those attending who wanted to continue the meetings for research purposes. A further and more subtle danger also arose in one or two cases. An uninterested majority coupled, perhaps, with some error of judgment on the part of the writer resulted occasionally in the assertion that the Boards of Studies were only doing the University's work for them. The extent of this feeling should not be exaggerated; nevertheless, it had been and still remains a cardinal principle that the research officer's position was that of adviser to the groups. The topics proposed for research and the direction of any which were taken up were always the responsibility of the group. In fact, a deal of guidance was generally necessary and, indeed, sought. Nevertheless, the principle of service rather than direction has always obtained and we would certainly not depart from it. As soon, therefore, as resistance was felt, it was clear that some other method of working was called for.

---

* A survey of some aspects of modern language teaching in certain secondary modern schools. *Educ. Rev.*, 5, 1953.

All experience to this point had suggested that whereas the formulation of a research project, the design of data gathering devices, the acquisition of samples and the collection of data could appropriately be carried out by the groups concerned, the analysis of the data and the first drafting of a research report were matters to be handled elsewhere. This was partly due to the fact that methods of statistical analysis were not known to most, to the fact that even if they had been known they would have made excessive demands on the time of lecturers and finally to the need, at some stage or other, for most research projects to come under the eye of a single person. With these needs to meet, therefore, and with the increasing administrative duties connected with the physical production of tests, questionnaires and so on, a secretarial assistant was appointed. Still more recently, a research assistant has been provided in addition so that three people are currently occupied with this work. It cannot be too strongly emphasised that if collective research work of the type with which this article is concerned, is to be carried out, a certain amount of technical and clerical assistance is essential. It is a waste of a University lecturer's time and talents to employ him in routine tasks of sorting and cyclostyling; he needs time also for his own reading and research and, if he is not to become too isolated, for a certain amount of University lecturing. It is the generous recognition of these needs by the University and the Institute of Education that has allowed such success as this collective work has achieved here.

III.—LATER RESEARCH ORGANISATION: THE TRAINING COLLEGES' RESEARCH GROUP.

When it became clear that the approach through Boards of Studies was not entirely successful, a new group was set up consisting only of those two or three lecturers from each training college or from each Board of Studies who were particularly keen to continue with research projects on a joint basis. This group, the Training Colleges' Research Group, is now in its sixth year. It has never met fewer than three times a year; often it has met six times a year and of late certain of its members have met even more frequently than this. The research officer and the research assistant have always attended all the meetings. The group has from the beginning conducted its business with a considerable degree of informality; it would have been perfectly possible to have made the procedures more formal by putting the research officer in the chair and initially there was, indeed, some pressure to do this This, however, would have been inconsistent with the policy of allowing the group to develop in its own way and to formulate its own policies and procedures. The group continued to work, therefore, without a chairman and without minutes for the first two years of its existence. It is the writer's opinion that any apparent loss of efficiency as a result of this was more than compensated for by the integration that the group developed. Since these first two years, however, as the work has become more varied and the group bigger (it now runs at about twenty-eight members per meeting) a chairman has been appointed from among its own members for a two-year tour of duty. The group has been singularly wise in its choice of chairmen; they, in turn, have worked without stint for the cause; by virtue of their office they are now members of the larger Research Committee of the Institute of Education, a body which meets annually to review the whole field of research under development within the Institute.

The topics which the group first chose for investigation were, naturally, those which affected its college activities most intimately. They were, therefore, more of an applied nature than pure. The group was early concerned and still remains so with the whole process of selection for admission to training colleges and particularly with the detection at selection or at an early stage of training of those who, for various reasons, were not likely to succeed. Non-academic reasons concerned with personality were of particular concern and a considerable amount of testing with personality schedules and related tests has been undertaken. The culmination of this work came with the testing of nearly a thousand students within the first two weeks of their College careers with what appeared from pilot experiments to be the best of the personality schedules and the comparison of these results with the judgments made on the students upon completion of their two-year training courses. A related piece of work considered, in the light of similar end-of-course criteria, the significance of much of the pre-college experience which was sometimes taken into account at selection interviews. Membership of various societies, the part played in school games, experience of work between leaving school and coming to college were typical of the activities and experiences which were evaluated in this way. This study was duly published*; the earlier personality testing study just mentioned is now being prepared for publication. While the need to publish results has not been over-stressed, the responsibility to do so has not been forgotten. This gives both a stimulus and a goal. Although it would be improper to dogmatise it is probable that occasional success in getting a paper accepted for publication interspersed with the distribution of locally produced cyclostyled or printed reports of all other researches, is an essential and integrating feature in the working of the group. It might well be that without these the group's activities would have petered out before this.

It is not the purpose of this article to give at length a description of the various studies which have been made. In the last three years there has been a move towards more fundamental research without neglecting the applied work rooted in experience with which the group was first concerned. Thus, there has been on the more fundamental side a large scale study, along Piaget lines, of children's concepts of area, the results of which are at present being analysed. Currently, the planning of a depth study of the quiet child is under way. On the more applied side, the problems of teachers in their probationary year are also being surveyed at the moment. There have also been studies which were quite small but dealt with matters requiring immediate attention, often for administrative or policy purposes, and studies which never got off the ground or never got anywhere even if they did.

IV.—PRESENT RESEARCH ORGANISATION.

As the group has increased in size and the projects have become more varied, it has recently become clear that there is a danger that the whole may become so unwieldy as to fragment, particularly since not all its members are now interested in all projects. There has always been a certain floating membership which changed with the topic under study and this feature has increased in the last two years. At the same time two other important matters of policy have come up for consideration; the first was whether teachers should become members of the group and the second was whether it should accept a greater responsibility for the planning and analysis of each research than it had hitherto. In principle, the first of these has been rejected, the group wishing to retain its identity as a training college group. Yet since much of its practical procedures have involved close co-operation with schools, a more formal procedure has now been adopted whereby head-teachers and class-room teachers are invited to planning conferences at an early stage in those researches in which schools are particularly involved. This is a new development but is interesting and promising. It is the more so since the writer has had rather little success in organising teacher groups for collective research similar to the lecturer groups. There are various reasons for this, not the least of which now is that this development would be big enough to demand yet another advisory appointment. In this Institute, this is shortly to be made. It is possible that these groups of teachers, invited at present to advise and co-operate in certain projects only, will themselves become nuclei for future research groups.

---

* The relationship between pre-college experience and teaching ability. *Educ. Rev.*, 13, 1960.

The present solution to the problem of increasing the group's responsibility for all stages of each project has been for the group to set up small steering committees of three of its members. These have the responsibility for carrying out the detailed work involved in any research and presenting their thoughts and plans to the bigger body at its terminal meetings. It is proposed that each new research should have its own steering committee ; currently, two are functioning. Although these steering committees have involved certain members of the main group in much more work and, although the standard and level of sophistication of the research is improving through their operation, it is by no means certain that their establishment is wholly without danger. If all the thinking and planning were to be done by them, leaving only a certain amount of rubber stamping to the larger group, there would be a distinct possibility of a collapse of the main body. But this is only a judgment of what might and not of what is happening.

V.—Recapitulation and conclusions.

It may appear that this article is unduly biographic. It is so because there are not so many similar groups in operation throughout the country and, not unnaturally, none of which the writer has such detailed experience. It must be stressed, if it is not already clear, that this group research work is only a small part of the total research effort of the Institute's Research Department. Much more goes on within the walls of the University than this article is concerned with. Yet this limited experiment here described has some unique features and has not been without success. The features which have contributed to such success are the rooting of the work in a department with previous research experience ; a generous University and Institute policy of staffing and clerical assistance, the seeking out of the few enthusiasts rather than the attempt to convert the many ; the giving of the group responsibility in and choice of its themes, its methods of working and its officers ; the maintenance of standards of rigour sufficient to allow the final research reports to be worthy of publication ; the anticipation of forces which might tend to weaken the integration of the group and, through periodic discussion devoted to underlying principles and through frequent visits to lecturers in their own colleges, to attempt to avoid them ; and, most of all, the willingness of the lecturers concerned to give their time to it.

As well as the principal effect of getting research carried out on a fairly large scale there are side effects which are wholly beneficial. More and more of our colleagues in training colleges have had some research experience and are the more appreciative of the value and the limitations of research. In turn, many of their students have, through the demands of the group, carried out a number of child studies on a relatively objective basis, a new and worthwhile experience for them. Finally, a number of the group's members have gone on to take higher degrees of University diplomas in education and, having done so, have returned to the group bringing a new level of expertise with them.

One cannot pretend that there are not other ways of working, that all, who ought to have become involved have done so or that there have not been failures and mistakes. It is the writer's conviction, however, that the Institute's original concept of its obligation to provide research facilities for its members was a correct one and his experience that other Institutes doing similar work share this conviction. It is also his conviction that Institutes of Education are, with their intimate links with the universities on the one hand and with the training colleges on the other, in a particularly favoured position to carry out work of the type described; it would almost be impossible to do so in any organisation which separated these two bodies. But whether they can continue to do so on a scale which the importance of such work justifies depends upon the provision of funds and facilities on a somewhat higher level than presently obtains.

(*Manuscript received* 22*nd March*, 1962.)

THE RELATION OF NEUROTICISM AND EXTRAVERSION TO INTELLIGENCE AND EDUCATIONAL ATTAINMENT

BY J. B. BIGGS

(*National Foundation for Educational Research*)

SUMMARY. Lynn and Gordon have recently reported an experiment the results of which are said to vindicate Eysenck's theory of conditionability. The present paper is an attempt to show (i) that, on theoretical and empirical grounds, Eysenck's theory is itself seriously open to question (ii) that, in any case, Lynn and Gordon's experiment appears to be founded upon a mistaken interpretation of Eysenckian theory, and (iii) that their experiment is liable to certain theoretical and methodological criticisms.

I.—INTRODUCTION.

LYNN and Gordon (1961) attempt to provide evidence for Eysenck's theory relating introversion and neuroticism to certain kinds of performance. Their paper raises several issues which one feels should be challenged. Their major claims are that introverts should take longer than extraverts on such a task as Raven's Progressive Matrices; that neurotics should be quicker than normals on the matrices; that neurotics should have good vocabulary scores, assuming that word learning is a simple conditioning process, and that there should be a curvilinear relationship between matrices performance and neuroticism. These predictions are said to be borne out by data presented in their paper.

An effective criticism of their paper, however, is possible only by first questioning their theoretical premises. These are based largely upon Eysenck's theory (1957) and to some extent upon drive theory (e.g., Taylor, 1956). The first section of the present paper, then, will be devoted to an examination of some aspects of Eysenckian theory. This will be followed by a critique of the way Lynn and Gordon interpret Eysenck, and finally, attention will be paid to some details of the experiment they report.

II.—EYSENCKIAN THEORY.

The fullest exposition of Eysenck's views in this field is to be found in *The Dynamics of Anxiety and Hysteria*, and accordingly, greatest emphasis will be placed on what he says there. In Hullian theory, performance is said to be a function of two things: drive and habit. Taylor (1956) ascribes differences in rate of conditioning to the former while Eysenck, on the contrary, rejects drive and suggests instead that we should look at habit.

Returning to Pavlov's original views on habit, Eysenck makes the now famous statement that " individuals in whom excitatory potentials are generated quickly and strongly are introverted in personality. Conversely, individuals who generate weak excitatory potentials slowly, and who generate strong inhibitory potentials quickly tend to be extraverted in personality " (p. 115). Conditioned reflexes should ,therefore, develop quickly and easily in introverts and slowly in extraverts—irrelevant drive, and Eysenck includes anxiety in this category, should have no effect on conditionability. Then comes the crucial experiment of Franks which is expounded at length on p. 115 and following.

Eysenck's interpretation of this experiment has been criticised elsewhere (e.g., Champion, 1961) and we need only to look at his treatment of the question of anxiety. In spite of the title of the book, Eysenck does not deal directly with the relationship between anxiety and performance: or rather, he subsumes it under the wider heading of introversion, and it must be assumed that the anxiety relationship is meant to be covered, almost by default, in the relationship of the build-up of inhibition to learning and performance. Inhibition build-up is said to be directly related to introversion* and to rate of conditionability. The point with regard to anxiety is seen more clearly, perhaps, in a later paper (1960), where Eysenck regards anxiety as " a conditioned fear response attached to previously neutral stimulus.'' In this paper, Eysenck is interested in immoral and criminal behaviour: extraverts are hard to condition, therefore, they do not experience anxiety (which is evidently regarded as being equivalent to guilt itself) so that there are no inward checks on anti-social or immoral behaviour.

Thus, if we take rate of conditionability as the basic factor, as he seems to do here, it follows that any two things which condition will be correlated. The introvert is said to condition easily while the psychopath does not. Since anxiety, moral behaviour, eyeblinks and presumably any form of conditioned learning are all conditionable, the introvert will manifest all these behaviours—assuming, of course, that he has been in the appropriate stimulus situations—and the extravert and psychopath will exhibit few of them. It would then follow that introverts should have higher anxiety scores than extraverts *and* a larger vocabulary (if, as Lynn and Gordon claim, vocabulary is acquired by conditioning) *and* manifest a better performance at mechanical, if not genuine problem, arithmetic—all these being concomitant, not causal, relationships. It is not quite playing fair to retort to this that neuroticism and introversion are along quite separate dimensions: if they are, then the model Eysenck offers can in no way replace that of Spence, but simply supplements it. Eysenck clearly states, however, that it is a question of replacing rather than of supplementing, and a moment's thought will show in fact that predictions based upon one theory would contradict those of the other.

Using Eysenck's theory strictly, then, one would predict direct linear functions between both introversion and anxiety and those kinds of learning and performance which rely upon conditioning to any extent. That such a positive linear relationship is rarely, if ever, found in practice, speaks for itself: it is certainly begging the question to test the predictions using linear statistical techniques such as correlation coefficients. If a relationship is suspected of being linear, and is, in fact, curvilinear, the r between the two variables will tend to be zero or very close to it. This is, in fact, what Lynn and Gordon find (see their rs for Neuroticism × Matrices (score) and Extraversion × Matrices (score)) and also what Eysenck consistently reports (see p.116 ff, 1957). Eysenck accepts curvilinearity as an empirical fact, then, and even goes so far as to put forward theories (the Taylor-Spence, and Mandler-Sarason (1952) models), one of which he is at pains to show is incompatible with his own, to account for it. His own theory, strictly applied, would not predict a curvilinear, but a linear, function—the higher the anxiety, the more introverted S tends to be, and therefore, the easier it is that conditioned learning takes place A study by

---

* Some of the physiological evidence that Eysenck cites for this is suspect. For instance he argues that introverts, in whom inhibition builds up slowly, should be able to tolerate depressant drugs to a greater degree than extraverts and he quotes the work of Shagass (1954) as confirming this. It should be noted, however, that Shagass's main findings were that tolerance of depressant drugs was highest in highly *anxious* patients, not introverts. Presumably, Eysenck would justify this by claiming that anxiety and introversion are correlated but this is hard to reconcile with the other point he continually stresses which is that anxiety and introversion are along quite separate dimensions.

Patterson, Helper and Wilcott (1960) shows that for relatively complex verbal conditioning, this is very far from being the case.

In sum, then, Eysenck's theory is internally inconsistent and only tenuously founded upon empirical fact. Shorn of all the trimimings, it is based upon rate of conditionability so that any two variables that are conditionable will be directly related to each other—introverted behaviour and vocabulary, or anxiety and rote learning, for example, would, on this theory, be positively correlated.

III.—LYNN AND GORDON'S PAPER.

Much of the above criticism of Eysenck's theory is directly applicable to Lynn and Gordon's paper, and there is no need to reiterate it. Further, than this, however, these two authors appear to be mistaken in their own interpretation of their theoretical background; it remains to discuss this, and to note some general criticisms of their paper.

Their adherence to *both* the Spence-Taylor model and the Eysenckian model is striking. Admittedly, they say that there is "no reason why both variables (i.e., drive and inhibition) should not be important," but, as they themselves found with their own data (cf. the *r*s of neuroticism and extraversion with vocabulary), and as Eysenck stresses (see page 107 in particular), the two accounts are incompatible. This fundamental confusion is perhaps seen more clearly in one of Lynn's earlier papers (1959).

He argues there that, according to Eysenckian theory, sustained work, such as is required of university students, demands two things: (i) strength of drive, and (ii) slow accumulation of reactive inhibition, i.e., introversion. Since Lynn clearly implies that neuroticism or drive has a direct causal relationship with academic performance—i.e., it evidently is not a side-prediction based upon the correlation between neuroticism and introversion—the first expectation that university students should exhibit a high degree of neuroticism, is presumably based upon drive theory, in an overly simplified form.

A central feature of both the 1959 and the 1961 papers, then, is that drive level facilitates academic and other forms of learning. His evidence for this suspect. First, he quotes Yates and Pidgeon's (1957) finding that at the age of 11, girls are better attainers than boys as being due to a higher drive level in girls. While it is probably true that, according to most acceptable measures of both drive and anxiety, girls do, in fact, manifest higher drive, over all age levels, than do boys, it is certainly not true that girls necessarily maintain their academic superiority after the age of 11. What usually happens—once sex biasses have been eliminated from the samples—is that differential attainment patterns emerge (Anastasi, 1958; Pidgeon, 1960), and these certainly cannot be explained on drive theory; *a fortiori* when an over-all and non discriminating criterion of performance, such as 'academic attainment' is used. The causation of sex differences in attainment is such a complex business that to account for it by a single factor, neuroticism, is a facile over-simplification, particularly when there are so many other personality and social factors which are also likely to be crucially important.

As further evidence for the beneficial effects of drive, Lynn quotes part of his own work (1955), showing positive correlations between anxiety and reading; he fails, however, to quote the remainder of this study reported later (1957) where he found an equally high but *negative* correlation between anxiety and arithmetic. Does he consider that reading and academic attainment in the university are less complex tasks than arithmetic? Not even the most ardent Hullian, however, could describe academic work as a simple task in which drive or anxiety could be facilitating—and if the multiplicative drive theory is admissable in this context, one further wonders how the other connection between rate of conditioning, introversion and conditioned fear (anxiety) can also apply. In other words, two quite separate models to explain how anxiety can affect performance are used, and they are each incompatible with the other. He does admit (1959) that " it seems that neuroticism has two different effects on attainment, a disorganising one on learning and performance in stress situations, but a facilitating one in so far as it motivates sustained work. Further at the educational levels of the university and school sixth form, its disorganising effects seem to be more than compensated for by its motivating powers. How far this is true at earlier educational levels seems to be a question well worth investigation."

There are several points to be considered here. It is clear, first of all, that there needs to be some very rigid definition of stress and neuroticism ; let us assume, as Lynn implies, that they are clearly distinct from each other. Many experimental studies into complex learning have used artificially induced stress (e.g., Cowen, 1952 ; Beier, 1951) while in others, the anxiety measure has been the usual questionnaire which Lynn would equate with neuroticism (e.g., Taylor and Spence, 1952 ; Farber and Spence, 1953 ; Spielberger, *et al.*, 1958). In both cases, the results have been, by and large, that complex learning is impaired by *both* stress and neuroticism. Where the ' complex learning ' has, in fact, been academic learning (or performance) it is not so easy to maintain the distinction since there have not been many studies investigating stress, as such. In one study, however, Faw (1956) found that examination performance of undergraduates under induced stress was unquestionably worsened. Examinations themselves can almost certainly be regarded as stress inducing in this sense ; especially in the case of those students who score highly on ' test ' anxiety questionnaires (see S. B. Sarason, *et al.*, 1960). All the evidence which Sarason marshals points very much to such stress being deleterious to performance at least in the elementary schoool—the one exception is the paper by Sarnoff, *et al.* (1960), mentioned by Lynn, where no anxiety effects were observed. The equation of ' test ' anxiety scores with neuroticism is by no means complete, but there is no doubt that the substantial correlation between these two scales is due largely to a common drive or anxiety factor ; or in other words, that ' general ' anxiety or neuroticism scales simply denote those individuals who are readily stressed in a wide variety of situations, and that ' test ' anxiety scales define a *particular* stress-provoking situation for experimental purposes. The distinction between neuroticism and stress thus appears partly to amount to a quantitative difference in the one underlying variable, anxiety, and partly to a matter of specificity, i.e., stress occurs whenever a given stimulus situation and a given individual reaction tendency come together. At any rate, the distinction is difficult to maintain from a practical or experimental point of view particularly when using the usual psychometric anxiety or neuroticism scales.

Anxiety at lower levels, before the stress or interfering effects operate, is unquestionably, as Lynn suggests, a motivating drive in that its reduction is reinforcing. This does not mean, however, that it necessarily acts as a drive at high levels, or in performance as opposed to learning. This last is an important distinction, for it is only in the performance function that anxiety is said to act as a drive by interacting *multiplicatively* with task complexity or already established response hierarchies : in learning, drive reduction, not drive multiplication, is the crucial concept and task complexity should not enter into drive interactions at all, except in so far as what are perceived as difficult tasks (which is not what learning theorists mean by task complexity) are themselves anxiety arousing. Lynn's own data indicate that anxiety probably does not act as a drive in the performance sense, i.e., by multiplyiug with task complexity, but this is not to say that mild, as opposed to stressful, anxiety will not motivate *learning*. On the contrary, it is clear that *any* external reward system, from merit cards to electrified grids, makes use of anxiety—fear of (not) obtaining the arbitrarily associated (positive)/negative reward—in a consistent and often effective way. Thus, neurotic students, who are easily aroused to specific anxieties, may at first show efficient learning, but these same students are the ones whose anxiety may readily become stressing (inhibiting) instead of motivating when external pressures become more insistent. There is thus no basic contradiction between the American and English studies as Lynn and Gordon suggest—it just depends upon the particular aspect of the anxious reaction and its correlates we happen to concentrate upon.

One does not need to search the literature very far to find examples of studies in which, time and time again, ' high drive,' neurotic or anxious children appear to compare less favourably than control groups in almost every aspect of school performance. Even if we take reading alone, which Lynn (1955) claims is facilitated by anxiety, this preceding statement appears to be true (Kerrick, 1955; Wirt and Broen, 1956; Lunzer, 1960). Although the last mentioned study used teacher-rated ' withdrawn,' ' aggressive ' and ' well-adjusted ' groups, the two former groups were found to exhibit more anxiety-type traits and obtained lower means scores in reading, arithmetic and matrices. However, Lunzer's findings could conceivably be due to the well-known halo-effect of teachers' ratings for adjustment (Wickman, 1928). Finally, McCandless and Castaneda (1956), whom Lynn refers to, clearly state that " there is a tendency for the more complicated skills (*reading*, arithmetic and concept performance) to suffer more interference from anxiety than mnemonic skills such as spelling." (italics not in original).

Lynn also quotes a finding of the present author (Biggs, 1959) as evidence that anxious children tend to be good readers. In this report, which was no more than a passing reference to a pilot study, it was stated that on a ' general ' anxiety scale a slight correlation was obtained between the scale score and scores on a standardised English test; more important, however, was the fact that ' number ' anxiety scores (i.e., where the child registered fear or worry specifically in arithmetical situations) produced a much higher positive correlation with English. This led to the conclusion that children, particularly girls, who were upset by their arithmetical prowess tended to compensate by increased effort in English. The general argument is similar to that of Levy (1943) who found that over-protected boys, for socially determined reasons, perform very well in verbal subjects.

In answer to Lynn's above question, then, it seems that at earlier educational levels, stress and, therefore, neuroticism or stress-tendency, is inimical to optimum performance, *in many school subjects*. The present writer is engaged in a project (to be published) investigating the relationship between anxiety, different aspects of arithmetical performance, and methods of teaching arithmetic, and preliminary results show that whether highly anxious subjects perform unexpectedly well or badly, depends upon many further factors, such as the mode of, and the methods of motivating, learning that are encouraged by different teaching methods, the success of such learning, the sex of subject and so forth. Generalizations in terms of one simple set of factors clearly seem to be inappropriate. It is highly likely that the intelligent child may be more able to deal adaptively with his anxiety than the average or unintelligent individual

—perhaps this is why sixth form and university subjects, both highly biassed samples, appear to be both more than usually anxious *and* good attainers. It is certainly unwise to say, as Lynn and Gordon do, that neuroticism is, therefore, one cause of academic attainment, and further, that the optimum level of neuroticism " appears to be in the region of half a standard deviation above the national average.''

The other matter arising out of the experimental and attainment studies is that of equating laboratory ' complex learning ' with academic learning. In order to be able to predict academic performance from theories which, in the main, have arisen out of laboratory studies, it is quite essential to establish from the start that academic learning embodies the same or analogous cognitive processes to those which are involved in the experimental tasks. Equating task difficulty with task complexity merely confuses the issue. Thus, while it is desirable and, at this stage, necessary to discover consistent interactions between academic performance and affective variables such as neuroticism; until some such kind of task analysis has been undertaken, it is certainly premature to apply a laboratory theory *in toto* to explain the numerous and largely unknown determinants of academic success.

We may conclude with the following observations on Lynn and Gordon's paper.

(1) In the absence of correlations between introversion and intelligence and neuroticism and intelligence, Lynn and Gordon conclude that explanations for the association between academic success and introversion and neuroticism cannot be due to an underlying interaction with intelligence. This conclusion, at least on the basis of the data we are given, is certainly not justified. Firstly, there is, in fact, a relationship between neuroticism and intelligence (as assessed by the matrices), albeit a curvilinear one, and secondly, the relationship between neuroticism and adacemic success was obtained on a sample not apparently identical with the present one. Obviously for the authors to reach the above conclusion, they can only do it by at least a two-way analysis of their data, on the one sample, and using statistical techniques which make no assumptions about the linearity of the variables analysed. Grice (1955) has shown that the Taylor Scale, and probably other similar scales, has significant interactions with intelligence and, therefore, all steps should be taken to ensure that this possibility has been eliminated—it certainly is not by the use of correlational techniques. In any case, since intelligence is not assessed by some kind of cerebral galvanometer but by performance on a set of tasks, surely a relationship between intelligence and both neuroticism and introversion would be exactly what one would expect on Eysenckian, Hullian or Lynnian grounds ?

(2) Having predicted that they may obtain at least one curvilinear function, they then base the majority of their conclusions upon linear correlations, thus precluding the possibility, without further analyses, of obtaining more evidence of curvilinearity—and to make confusion worse confounded, attempt a greater degree of refinement by partial correlations. Again, the only legitimate way to analyse data of this sort, the statistical properties of which are dubious, is by discrete two, three or even four-way groupings ; and if the sample is not large enough to permit this, then it is not permissible to draw the conclusions which Lynn does.

(3) The conclusion that it is in the later stages of a task that introverts become slower than extraverts is surely a *non sequitur* from the proposition that extraverts build up inhibition much more quickly and easily than introverts ? Since Lynn and Gordon admit that both extraverts and introverts attempted the same number of items, then the (very small) correlation

with matrices time and introversion is a contradiction of their theory, not a vindication of it. One must regret once again that sub-group means and S.Ds. were not presented in place of or in addition to their extremely doubtful correlation table.

(4) It will be apparent by now that their finding of the curvilinearity of neuroticism against matrices score, while not unexpected, is awkward for them to explain. Their language is a little extravagant when it is claimed that "it confirms the theory that for this type of task there is a golden mean of neuroticism or anxiety . . ." The Yerkes-Dodson law, which is the theory they claim to be confirming, is not, in fact, a theory, but an empirical statement of fact, and it was not originally proclaimed as being suitable for 'this kind of task,' i.e., Matrices, or even, for that matter, has it to do with anxiety as such, but motivation in general. The law was first derived from animal experiments where some form of conditioning was used: its application to the quite different *kind* of learning required by the Progressive Matrices requires more than Lynn and Gordon's bland assurance that the matrices "may be regarded as a task of moderate complexity (presumably they mean difficulty) for university students." As already indicated, complexity is a very different, although related, matter from difficulty—when complex S-R learning is virtually equated with moderately complex insightful learning, the result is psychological nonsense exactly on a par with their implied contention that academic success is a simple task where drive can be facilitatory,. However, the inadequacy of their analysis of the structure of different tasks, and the psychological processes thereby involved, has been mentioned already. Without such an analysis, their prediction from an empirical, as opposed to a theoretical, statement is simply a waste of time, particularly when the matrices test was not analysed according to item complexity—which is surely the whole point of discussing the Yerkes-Dodson law at all. Having done this, there would be a great deal to be said for trying to explain their obtained curvilinearity—or the Yerkes-Dodson law itself—in theoretical terms, but if the criticisms suggested in the present paper are at all valid, it seems that this would be a difficult task indeed.

IV.—REFERENCES.

ANASTASI, A. (1958). *Differential Psychology*. Macmillan, New York.

BEIER, E. G. (1951). The effect of induced anxiety upon the flexibility of intellectual functioning. *Psych. Monogr.*

BIGGS, J. B. (1959). Teaching of mathematics II. Attitudes to arithmetic-number anxiety. *Ed. Res.*, Vol. 1, No. 3, 6-21.

CHAMPION, R. A. (1961). Some comments on Eysenck's treatment of modern learning theory. *Brit. J. Psychol.*, 52, 167-173.

COWEN, E. L. (1952). The influence of varying degrees of psychological stress on problem-solving rigidity. *J. Ab. Soc. Psychol.*, 47, 512-9.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.

EYSENCK, H. J. (1960). Symposium. The development of moral values in children. VII—The contribution of learning theory. *Brit. J. Educ. Psychol.*, 30, 11-21.

FARBER, I., and SPENCE, K. (1953). Complex learning and conditioning as a functioning of anxiety. *J. Exp. Psychol.*, 45, 120-125.

FAW, V. E. (1957). Learning to deal with stress situations. *J. Educ. Psychol.*, 48, 134-144.

GRICE, G. (1955). Discrimination reaction time as function of anxiety and intelligence. *J. Abnorm. and Soc. Psychol.*, 50, 71-74.

KERRICK, J. S. (1955). Some correlates of the Taylor manifest anxiety scale. *J. Abnorm. and Soc. Psychol.*, 50, 75-77.

LEVY, D. M. (1943). *Maternal Over-Protection.* New York, Columbia Univ. Press.

LUNZER, E. A. (1960). Aggressive and withdrawing children in the normal school. I—Patterns of behaviour. II—Disparity in Attainment. *Brit. J. Educ. Psychol.,* 30, 1-10, 119-123.

LYNN, R. (1955). Personality factors in reading achievement. *Proceedings of the Royal Society of Medicine,* Vol. 48, 996-998.

LYNN, R. (1957). Temperamental characteristics related to disparity of attainment in reading and arithmetic. *Brit. J. Educ. Psychol.,* 27, 62-67.

LYNN, R. (1959). Two personality characteristics related to academic achievement. *Brit. J. Educ. Psychol.,* 29, 231-216.

LYNN, R., and GORDON, I. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.,* 31, 194-203.

MANDLER, G., and SARASON, S. B. (1952). A study of anxiety and learning. *J. Abnorm. and Soc. Psychol.,* 47, 166-173.

MCANDLESS, J., and CASTANEDA, A. (1956). Anxiety in children, school achievement and intelligence. *Child Dev.,* 27, 379-382.

PATTERSON, G. R., HELPER, M. E., and WILCOTT, R. C. (1960). Anxiety and verbal conditioning in children. *Child Dev.,* 31, 101-108.

PIGEON, D. A. (1960). A national survey of the ability and attainment of children at three age levels. *Brit. J. Educ. Psychol.,* 30, 124-133.

SARASON, S. B., *et al.* (1960). *Anxiety in Elementary School Children.* New York: John Wiley and Sons, Inc.

SARNOFF, I., and SARASON, S. B., *et al.* (1959). Test anxiety in the the 11 plus examination *Brit. J. Educ. Psychol.,* 29, 9-16.

SHAGASS, G. (1954). The sedation threshold. *EEG. Clin. Neurophysiol.,* 6, 221-5.

SPIELBERGER, C. D., *et al.* (1958). Complex incidental learning as function of anxiety and task difficulty. *J. Exp. Psychol.,* 56, 58-61.

TAYLOR, J. A. (1956). Drive theory and manifest anxiety. *Psychol. Bull.* 53, 303-320.

TAYLOR, J. A., and SPENCE, K. W. (1952). The relationship of anxiety level to performance in serial learning. *J. Exp. Psychol.,* 44, 61-64.

WICKMAN, E. K. (1928). *Children's Behaviour and Teacher's Attitudes.* New York: The Commonwealth Fund.

WIRT, R. D., and BROEN, W. E. (1956). The relationsip of the C.M.A.S. to the concept of anxiety as used in the clinic. *J. Consult. Psychol.,* 20, 482.

YATES, A., and PIDGEON, D. A. (1957). *Admission to Grammar Schools.* London: Newnes.

COMMENTS ON THE ARTICLE BY J. B. BIGGS

By R. LYNN
(*University of Exeter*)

Biggs has raised a number of interesting questions, which I shall try to deal with in turn. His first points are concerned with the adequacy of Eysenck's personality theory, and his argument here is that Eysenck's theory is so internally inconsistent and so tenuously founded on fact that it is a waste of time to try to apply it to educational problems. Biggs' first attack on the internal consistency of Eysenck's theory is concerned with the Yerkes-Dodson law : he argues that Eysenck's theory implies a linear relation between anxiety and speed of conditioning, whereas Eysenck erroneously attempts to deduce the curvilinear relationship. Biggs's argument here rests on the logical error of supposing that when A (anxiety) correlates with B (introversion), and B correlates with C (conditionability), it follows that A must correlate with C. This is, of course, a false inference and Biggs's attack on Eysenck on this score must be dismissed.

A second inconsistency in Eysenck's theory, according to Biggs, concerns the relation between anxiety and introversion. On the one hand, Biggs asserts, Eysenck maintains anxiety is associated with introversion, but on the other hand ' he continually stresses that anxiety and introversion are along quite separate dimensions." I was extremely surprised to read Bigsg's assertion on this question, since the association of anxiety with both introversion and neuroticism is one of the most fundamental notions of Eysenck's theory. As Biggs does not give any reference it is hard to see how he can have got hold of the idea that Eysenck ' continually stresses ' the independence of these two characteristics, but it would seem that he has misunderstood Eysenck's theory on this point

Biggs also criticises Eysenck's theory for being tenuously founded on empirical fact and raises the issue of the relationship of conditionability to introversion and anxiety, concluding that " a positive linear relationship is rarely, if ever, found." Once again Biggs gives no references to support his assertions, but I should have thought that the weight of the evidence is now clearly in favour of a relationship between anxiety and conditionability in simple tasks (a considerable amount of evidence is reviewed by Taylor (1956) and Spence (1960)). There is less evidence on the relationship between conditionability and introversion, but four studies reviewed by Willett (1960) all show some association in the direction predicted by Eysenck's theory. A fifth study reported significant introversion-extraversion differences in verbal conditioning (Eysenck, 1959a) It is true that some of the correlations have been very low and some doubts might reasonably be expressed about the extent of the introversion-conditionability association, but to state that the relation is " rarely if ever found " is certainly unjustified.

After his attempt to demolish Eysenck's theory, Biggs turns his attention to my application of the theory to the question of educational attainment. His first point here is concerned with my adherence to both the Eysenck (introversion and reactive inhibition) and the Spence (anxiety and drive) theories of conditioning. Biggs holds that these two theories are incompatible and further that Eysenck also regards them as incompatible. Unfortunately, Biggs does not explain why he thinks the two theories are incompatible but merely asserts that the incompatability will be evident with " a moment's thought." I believe that Biggs is quite wrong about this. In the first place, it should be obvious that where one variable A (introversion) is associated with another variable B (conditionability), no inference at all can be made about the relation of B to another variable C (drive, neuroticism or anxiety). It is also hard to see how Biggs can think that Eysenck follows this extraordinary logic. Eysenck's (1957, chapt. 4) position on this question seems quite plain. He first suggests the hypothesis that introversion should correlate with conditionability and compares it with Spence's theory that drive correlates with conditionability. He then cites Frank's experiment which shows that normals condition more quickly than hysterics, which is counter to the Spence theory, but supports his own. However, this finding (taken by itself) only shows that introversion is a more important factor in conditionability than neuroticism or anxiety because hysterics might still condition more quickly than normal subjects matched for introversion (i.e., stable extraverts). If this were so, then both the Eysenck and the Spence theories would be supported. Eysenck clearly regards the validity of the Spence theory as still an open question and a matter of fact to be investigated, and he goes on to cite a result showing no significant correlation between neuroticism and, conditionability. It is, therefore, clear that Eysenck regards the correctness of his own theory and incorrectness of Spence's as matters of fact, and is in no way arguing that the two are logically incompatible.

While Eysenck and I are in agreement that there is no logical incompatibility between the introversion-reactive inhibition and anxiety-drive theories of conditionability, my own position on the empirical justification for Spence's theory differs somewhat from Eysenck's. This difference is possibly due to the new evidence in favour of Spence's theory which has accumulated since Eysenck's 1956 book. However, that may be, my own view is that Frank's finding of a negligible correlation between neuroticism and conditionability is outweighed by the great variety of evidence offered by Spence in favour of a correlation between drive and conditionability, especially the experiments showing a positive relationship between condionability and (*a*) UCS intensity, (*b*) the intensity of interspersed irrelevant stimuli, and (*c*) physiological measures of drive (Spence, 1960). Consequently, taking the evidence as a whole, I favoured the view that conditionability is related to both introversion and drive.

In his discussion of the relation of anxiety to attainment, Biggs seems to be agreeing with my own hypothesis that the relationship obeys the curvilinear Yerkes-Dodson principle, i.e., some moderate level of anxiety is optimum for efficient performance. Biggs's interpretation of his own positive correlations between various measures of anxiety and reading follows my own hypothesis that anxiety can act as a drive motivating attainment. I have myself pointed out that very high levels of anxiety impair academic attainment (Lynn, 1955) and that many American investigators have reported negative correlations between anxiety and attainment (Lynn and Gordon, 1961), I did not discuss these at any length, but since Biggs argues that these American findings (e.g., McCandless and Cartananda, 1956; Sarason, *et al*, 1960) are strong evidence against my hypothesis that anxiety can motivate attainment (although Biggs himself uses my hypothesis to explain his own results), I should like to raise the question of the validity of these tests. Most of the American investigations have used group questionnaire tests to measure anxiety. When I began investigating this question, I tried this method of administering group questionnaires and I found that, no matter how much one gave the usual instructions about there being no right or wrong answers, that the truth was wanted, that the teacher would not be shown the results, etc., many children put down the same answers as their neighbour and several would ask afterwards how many they had got right. Another disturbing factor in the use of group questionnaires is that when lie scales are used quite high lie scale scores are obtained (Sarason, *et al.*, 1960). For these reasons I came to doubt the validity of group questionnaires given to children and tried to assess anxiety by getting to know the children individually and asking standard questions in an informal interview situation. My doubts about group questionnaires for children were later strengthened by the finding of Thorpe and James (1957) that children attending an out-patient department for emotional difficulties scored *lower* than matched controls on a questionnaire of neuroticism administered in school! About the same time, Wirt and Broen (1956) reported an attempt to validate the Children's Manifest Anxiety Scale (CMAS), which is the test used by Cartananda in his investigations showing negative correlations between anxiety and attainment. Wirt and Broen found that the CMAS did not differentiate clinic children from normals, nor did it differentiate children rated as anxious by psychologists and teachers. These investigators concluded that the anxiety questionnaire does not measure anxiety. In view of these findings, the American results reporting negative correlations between anxiety and attainment, and also that reported recently in England by Hallwood (1961), ought to be viewed with considerable reservation. It was because of these difficulties that I regarded the relationship of anxiety to attainment in school children as still an open one, and also that I raised the question of sex differences in anxiety and attainment, since it is agreed by many (including Biggs) that girls are more anxious than boys and better attainers at the age of 11 This appears to be to me one of the few facts available in this field which is so lacking in valid tests. But I would not agree with Biggs that there is any clear evidence that girls lose this lead at later ages ; the complications of early school leaving and independent schools make sex differences at later ages an extremely difficult matter to determine and I doubt whether acceptable evidence is available I have never argued, as Biggs suggests, that girls' greater neuroticism is the *only* factor accounting for their superior attainment and it would obviously be foolish to do so, since additional unsuspected factors are always liable to be discovered.

It is puzzling that, while Biggs appears to agree with my hypothesis that the Yerkes-Dodson principle operates in educational attainment, he proceeds to take me to task for attempting to apply this principle to the educational fields. His grounds for this objection are that the Yerkes-Dodson law was first established on animals and cannot legitimately be applied to human problem solving task. As far as the factual side of this objection is concerned, the Yerkes-Dodson law has been demonstrated in human beings on a very wide variety of tasks (the most recent review is that of Broen and Storms (1961)). In attempting to apply this law to the question of academic attainment I was following an established procedure in science by which the basic laws are generally discovered in strict laboratory conditions and are then applied to more everyday circumstances. It is, of course, true that allowance must be made for complicating factors However, the best method of determining how far laboratory principles can usefully be employed in an applied field such as educational psychology is surely to try them out rather than argue *a priori* whether the situations are or are not comparable. One advantage of attempting to apply laboratory principles to applied fields is that it makes one look for relationships which otherwise escape notice: this well-known principle is exemplified quite well in this particular case, where most investigators have reported their results in terms of correlations without apparently looking for a curvilinear relationship.

Some of Biggs's later points can be dealt with more briefly. The balance of the evidence seems to me to be that there is a curvilinear relation between anxiety and both learning efficiency and performance (Broen and Storms, 1961). Many findings showing a positive relation between simple conditioning and anxiety have been reported by Spence (1960). I have myself stressed the importance of looking for curvilinear relationships and the data were examined for these, but the only significant departure from linearity was that reported. I partly agree with Biggs in that from the general theory two results might have been expected about the time taken by introverts and extraverts on the later stages of the task : as inhibition develops, it might have the effect of slowing down all subjects, in which case extraverts would be slower, but it might be that extraverts become more slapdash (partly because they are less socialised in general) while introverts continue to be painstaking. There is nothing objectionable about a theory yielding ambiguous predictions on a few issues as long as it does not do it too often. The actual finding that extraverts get quicker seems explicable by the thoery and is in accordance with a number of other findings. Biggs does not define the distinction he makes between complexity and difficulty, but I use the terms to denote a task in which the correct habit is not dominant. I agree that it is hard to determine the degree of difficulty or complexity of tasks like learning to read, do arithmetic, do the matrices or follow a university course. Nevertheless, optimum performance on almost any task has a curvilinear relationship with anxiety (Broen and Storms, 1961) and hence it seems reasonable to look for the curvilinear relationship in academic performance. Biggs' statement that my 1957 paper reported a negative correlation between anxiety and attainment in arithmetic is a misunderstanding. The paper reported that anxiety is related to attainment in reading being better than attainment in arithmetic and, in fact, there was a low positive correlation between anxiety and attainment in arithmetic.

Finally, I think that Biggs and I have much more in common than the differences discussed in these papers might suggest. It is evident from Biggs's researches and his thoughtful comments on my own work that we are agreed about the importance of personality factors in attainment, and especially the question of anxiety I welcome the fact that Biggs is doing research in this field and I shall read his results with great interest.

REFERENCES.

BROEN, W. E., and STORMS, L. H. (1961). A reaction potential ceiling and response decrements in complex situations. *Psychol., Rev.*, 68, 405-415.

EYSENCK, H. J. (1959a). Personality and verbal conditioning. *Psychol. Reps.*, 5, 570.

EYSENCK, H. J. (1959b). *Manual of the Maudsley Personality Inventory*. Univ. of Lond. Press.

HALLWORTH, H. J. (1961). Anxiety in secondary modern and grammar school children. *Brit. J. educ. Psychol.*, 31, 281-291.

LYNN, R. (1955). Personality factors in reading achievement. *Proc. Roy. Soc. Med.*, 48, 996-998.

LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.

MCCANDLESS, B R., and CARTANANDA, A. (1956). Anxiety in children, school achievement, and intelligence. *Child Development*, 27, 379-382.

SARASON, S. B., *et al.* (1960). *Anxiety in Elementary School Children*. London: John Willey.

SPENCE, K. W. (1960). *Behaviour Theory and Learning*. Englewood Cliffs · Prentice-Hall.

TAYLOR, J. A. (1956). Drive, theory and manifest anxiety *Psychol. Bull.* 53, 303-320.

THORPE, J. G., and JAMES, D. P.(1957). Neuroticism in children. I—An investigaton of normal and neurotic group differences. *Brit. J. Psychol.*, 48, 26-34.

WILLETT, R. A. (1960). Measures of learning and conditioning. In *Experiments in Personality*, ed. H. J. Eysenck. London · Routledge and Kegan Paul.

WIRT, R. D., and BROEN, W. E. (1956). The relation of the children's manifest anxiety scale to the concept of anxiety as used in the clinic. *J. Consult. Psychol.*, 20, 482.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

The Relation of Social Adaptation, Emotional Adjustment and Moral Judgment to Intelligence in Primary School Children.

By W. K. BRENNAN

(Abstract of thesis submitted for degree of M.Ed. in the University of Manchester.)

OBJECT.

The main aim of the research was to study the relation of social adaptation to different levels of intelligence and to age. The relation of social adaptation to moral judgment and the relation of both to emotional adjustment, were also studied.

METHOD.

The Cornwell Orally Presented Group Test of Intelligence for Juniors (Cornwell, 1950) was administered to 548 primary school children in Farnworth, Lancashire. From these, seven experimental groups of twenty children were selected, all matched for size of family and occupation of father, but with controlled contrasts in age and intelligence as shown in Table 1.

TABLE 1

MEAN CHRONOLOGICAL AND MENTAL AGES (MONTHS) IN THE EXPERIMENTAL GROUPS n=20.

		Chronological Age		
		11	9½	8
Mental Age	11	Average C.A. 132·70 M.A. 132·65	Bright C.A. 115·85 M.A. 134·85	— —
	9½	Dull C.A. 134·10 M.A. 111·20	Average C.A. 116·15 M.A. 114·00	Bright C.A. 98·00 M.A. 112·85
	8	—	Dull C.A. 117·00 M.A. 97·25	Average C.A. 98·05 M.A. 97·30

The Vineland-Manchester Scales of Social Adaptation (Lunzer, 1959) and a test of moral judgment based on the work of J. Piaget (1932) were administered to each child individually, and for each child teachers completed a schedule of the Bristol Social Adjustment Guides (Stott, 1958) as a measure of emotional adjustment.

A scalogram technique (Guttman, 1950) was used to establish the scalability of the moral judgment questions; total scores from questions forming the scale are used in later analysis.

Two separate analyses of variance are applied to scores from the Vineland-Manchester and moral judgment tests. Each covers the results from four of the seven experimental groups: the first stage of the analysis relates to groups 8 Bright, 8 Average, 9½ Average and 9½ Dull; the second to groups 9½ Bright, 9½ Average, 11 Average and 11 Dull. The 9½ Average group is common to both analyses. In effect an attempt is being made to answer the following questions:

(1) In the first analysis :
(*a*) Is the 8-year-old child who is bright more like an average 8 than a average 9½-year-old ?
(*b*) Conversely, is the 9½ child who is dull more like the average 8-year-old ?

(2) In the second analysis :
(*a*) Is the 9½-year-old who is bright more like the average 11 year-old ?
(*b*) Conversely, is the 11-year-old who is dull more like the average 9½-year-old?

Data at each age-level are re-arranged into emotionally well, moderately or poorly adjusted sub-groups on the basis of the children's Bristol Guide scores. Analysis of co-variance of intelligence, first with Vineland-Manchester and then with moral judgment scores, allows the adjustment of these scores to a common initial mean for intelligence score. Following significant F ratios, the 'adjusted means' are calculated and differences between them tested by *t* scores. The technique is also used to analyse the Social Adaptation scores of sub-groups with high, moderate or low moral judgment scores.

RESULTS.

(1) *Social Adaptation.*
(*a*) *Social Responsbility Scores.*

TABLE 2

MEAN SOCIAL RESPONSIBILITY SCORES.

		Chronological Age		
		11	9½	8
Mental Age	11	Average 45·00	Bright 40·15	—
	9½	Dull 42·40	Average 37·40	Bright 31·05
	8	—	Dull 35·85	Average 30·00

Both analyses of variance show C.A. as a significant factor at the ·001 level. M.A. is significant only in the 9½–11 analysis ($P < \cdot 05$).

(*b*) *Social Perspective Scores.*

TABLE 3

MEAN SOCIAL PERSPECTIVE SCORES.

		Chronological Age		
		11	9½	8
Mental Age	11	Average 21·15	Bright 23·35	—
	9½	Dull 14·35	Average 15·05	Bright 13·05
	8	—	Dull 9·70	Average 7·45

Both analyses of variance show M A. as the only factor attaining a high degree of significance ($P < \cdot 01$). The variance attributable to C.A does not reach the ·05 level.

(*c*) *Relation between Emotional Adjustment Level and Social Adaptation Scores.*

In this section differences in mean intelligence score consistently favour the sub-groups with better emotional adjustment, though differences are significant only at the 9½-year-level ($P < \cdot 01$).

(1) *Social Responsibility Scores.*

TABLE 4

MEAN SOCIAL RESPONSIBLITY SCORES IN EMOTIONAL ADJUSTMENT SUB-GROUPS.

Age	Well Adjusted	Moderately Adjusted	Poorly Adjusted
8	28·06	29·91	34·58
9½	38·30	38·55	36·62
11	43·25	43·67	44·10

Between-groups F ratios from analysis of variance did not reach the ·05 level of significance. Analysis of co-variance, however, showed between-groups F ratio at the 8-year level as significant ($P < \cdot 05$). Means adjusted for intelligence were 27·73, 30·03 and 34·94. It is of some interest to note that the children who are less well-adjusted appear to be carrying a higher burden of social responsibility.

(2) *Social Perspective Scores.*

TABLE 5

MEAN SOCIAL PERSPECTIVE SCORES IN EMOTIONAL ADJUSTMENT SUB-GROUPS.

Age	Well Adjusted	Moderately Adjusted	Poorly Adjusted
8	11·82	8·63	4·03
9½	22·75	15·37	10·72
11	20·00	18·89	15·74

The only significant differences revealed by analysis of variance were at the 9½-year level ($P < \cdot 01$). But the marked differences between these sub-groups are significant only if differences in intelligence are not taken into account The significance disappears following analysis of co-variance which adjusts perspective scores to a common initial mean for intelligence ($P > \cdot 10$).

(2) *Moral Judgment*

(*a*) *Scalogram Analysis.*

(1) The ten moral judgment questions were not ' scalable,' (ratio of reproducibility=0·80) ; three questions with ' grouped non-scale errors ' were eliminated.

(2) The remaining seven questions are scalable (ratio of reproducibility=0·92 ; minimal marginal reproducibility=0·69), and the totals from these provide data for analysis.

(*b*) *Relation between M.A., C.A. and Moral Judgment Scores.*

TABLE 6

MEAN MORAL JUDGMENT SCORES.

		Chronological Age		
		11	9½	8
Mental Age	11	Average 6·10	Bright 6·10	—
	9½	Dull 5·45	Average 5·15	Bright 3·50
	8	—	Dull 4·05	Average 1·85

For the 8–9½-year groups both C.A. and M.A. are significant, $(P < \cdot 001)$. For the 9½–11 groups only M.A. is significant $(P < \cdot 05)$.

(*c*) *Relation between Emotional Adjustment and Moral Judgment Scores.*

TABLE 7

MEAN MORAL JUDGMENT SCORES IN EMOTIONAL ADJUSTMENT SUB-GROUPS.

Age	Well Adjusted	Moderately Adjusted	Poorly Adjusted
8	3·29	2·18	2·25
9½	6·10	5·00	4·24
11	6·42	5·78	5·37

The only significant between-group F ratio is on the variance at 9½ years $(P < \cdot 05)$; interpreted for heterogeneity of variance). However, with moral judgment scores adjusted to a common initial mean for intelligence score by analysis of co-variance the F ratio is not significant.

(*d*) *Relation between Moral Judgment and Social Adaptation Scores.*

(1) *Social Responsibility.*

The analysis does not reveal any consistent or significant relationships between moral judgment and social responsibility.

(2) *Social Perspective.*

TABLE 8

MEAN SOCIAL PERSPECTIVE SCORES IN MORAL JUDGMENT SUB-GROUPS.

Age	High M.J.	Moderate M.J.	Low M.J.
8	12·00	11·12	8·45
9½	17·41	16·80	11·87
11	18·64	18·11	14·23

These figures show a consistent association between high moral judgment and high perspective scores, with F significant only at 9½ years ($P < \cdot 025$; interpreted for heterogeneity of variance). Differences between low and high sub-group means ($t < \cdot 02$) and low and moderate ($t < \cdot 05$) are statistically significant.

CONCLUSIONS.

Within the control limits of the research, the following conclusions are formulated:

(1) Perspective, the cognitive aspect of social adaptation, is positively associated with mental age.

(2) Responsibility, the executive aspect, is positively associated with chronological age.

(3) Neither perspective nor responsibility is significantly related to emotional adjustment level.

(4) Seven moral judgment questions are scalable and increasing total score indicates movement away from 'moral realism' as defined by Piaget (1932).

(5) Both C.A. and M.A. are significantly related to moral judgment up to 9½ years. From 9½–11 years, the significance of N.A. is reduced and C.A. becomes insignificant.

(6) High perspective score is associated with low 'moral realism,' though the association is significant only at 9½ years.

(7) Where age and intelligence do not differ significantly, moral judgment is not significantly associated with responsibility or emotional adjustment.

(8) There is a consistent and positive association between mean intelligence score and level of emotional adjustment, differences in intelligence favouring better adjusted sub-groups. The association is significant only at the 9½ year level.

SELECTED BIBLIOGRAPHY.

CORNWELL, J. (1950). *The Construction, Standardisation and Validation of an Orally Presented Group Test of Intelligence for Children between the Ages of* 8 *and* 11. Unpublished M.A. Thesis. University of Birmingham.

GUTTMAN, L. (1950). Chapters 2 and 3 in *Measurement and Predictions.* Ed., Stouffer, S. A. London: Oxford University Press.

LUNZER, E. A. (1959). Introduction to the *Vineland-Manchester Scales of Social Adaptation.* Unpublished cycolstyle. University of Manchester.

—— (1960). Aggressive and withdrawing children in the primary school. *Brit. J. educ. Psychol.*, 30, 1-10, 119-123.

PIAGET, J. (1932). *The Moral Judgment of the Child.* London: Kegan Paul.

STOTT, D. H. (1958). *The Social Adjustment of Children. Manual to the Bristol Social Adjustment Guides.* London: Univ. of London Press.

An Examination of the Causes of Premature and Early Leaving in Selected Grammar Schools of South Wales.

BY H. G. EVANS

(Abstract of thesis submitted for the degree of M.A. in the University of Wales, University College of Swansea, 1961)

BACKGROUND.

This thesis attempts to examine some of the reasons why many of the 256 pupils who went to three South Wales grammar schools in September, 1953, left those schools prematurely. The three schools chosen are all situated relatively close to one another in the anthracite area of the coalfield and have very similar geographical and sociological environments. The large majority of the families in the area are Welsh speaking. They are also mostly Nonconformist in their religion and are nearly wholly dependent on the coal mining industry for a living.

PROCEDURE.

Twelve separate questionnaires were drawn up and used in the investigation in order to obtain as comprehensive a background picture as possible for each of the 256 pupils—both in school and at home. Most of the pupils filled in questionnaires while they were still at school and so did their parents, teachers and headmasters. The small number of premature leavers who had already left were interviewed by the investigator in their homes, and so were all the other pupils who subsequently left prematurely or early.

Statistical tables were then drawn up from the material contained in the thousand odd questionnaires that were completed, the tables being set out in such a way as to allow easy comparison between the premature and early leavers on the one hand and the sixth form entrants on the other. The Chi squared test was used to test the statistical significance of the differences.

RESULTS.

The rate of premature and early leaving from the three schools during the period studied was found to be appreciably lower than the rate worked out for England as a whole by the Central Advisory Council for Education (England) in their " Early Leaving " report. Fewer of the leavers were also thought capable of doing advanced work. Wastage was still extensive, however, particularly amongst the girls. Premature leaving was considered to be more serious than early leaving in that far more of the premature than the early leavers left against the advice of their headmasters and entered poor, badly chosen occupations with no future prospects.

In examining the influence of the home the thesis comes to the conclusion that there is a strong relationship between the environment of the home—both material and cultural—and length of grammar school life. The statistical tables show clearly that far more of the parents of the sixth form entrants than those of the premature and early leavers were working in non-manual occupations and had themselves been educated at grammar school. More of the parents of the sixth form entrants than those of the premature and early leavers lived in what could be termed better-type houses, e.g , in private detached and semi-detached houses. The families of the sixth form entrants were smaller, the density of persons per room in their homes was lower and there was better provision for homework.

The parents of the sixth form entrants also showed more interest in the education of their children, tended to speak Welsh in their homes and be Nonconformist in their religion more than the parents of the premature and early leavers. No large difference was seen, however, between the two groups with regard to family income and possessions.

In examining the influence of the pupils themselves, the tables show the premature and early leavers to be poorer grammar school pupils than the sixth form entrants. They did worse in the entrance examinations, in the end of the year school terminals and in the G.C.E. Ordinary level examinations. They had sat the entrance examination a number of times, were older on entering grammar school and were absent more often. They found school work more difficult, were less interested in it, and spent less time on their homework. They were not so happy at school, read less, belonged to fewer school and non-school clubs and societies, but visited the cinema more often. Their teachers also gave them poorer temperamental trait assessments.

No correlation was observed between the proportion of the age group entering a grammar school and the proportion leaving prematurely or early ; the school with by far the highest rate of intake had by far the lowest rate of wastage. Pupils from certain primary schools, however, succeeded far better at grammar school than pupils from certain others, while pupils transferred from secondary modern schools very largely failed to have long grammar school careers. The grammar schools themselves, also probably helped to influence their own rates of wastage, the school which restricted sixth form entry most rigidly having the highest rate of wastage.

Local employment openings played little part in early school-leaving, there being a general lack of such openings in the area studied. The fact that many of the pupils were reaching, or had reached, their academic peak was probably a far more important factor behind leaving. Boredom and loss of interest, brought about largely because school work was becoming too difficult, was quoted as the most powerful reason for leaving by the premature leavers themselves and by their parents and headmasters. Failure, or that they had done badly, in the G.C.E. Ordinary examinations was quoted also as an important reason for leaving by many of the early leavers.

BOOK REVIEWS

ARGYLE, MICHAEL, and TIZZARD, JACK (Edits.). *The British Journal of Social and Clinical Psychology*. To be issued in parts, three times per year. 50s. per Volume (post free). Separate parts 20s. (Subscription rates for members of the British Psychological Society, 20s. per Volume.)

The first part of a new Journal of the British Psychological Society appeared in February this year. Called *The British Journal of Social and Clinical Psychology*, its coverage, as is indicated by its title, may at times overlap slightly with this *Journal*. There is, for example, an article in this first number by Oliver and Butcher on *Teachers' Attitudes to Education*. There is no doubt, however, that for some time there has been a considerable need for a place to publish British papers of an experimental kind in the fields of abnormal and clinical psychology and in many aspects of social psychology.

We wish the new Journal well and commend it to our readers for general interest and for relevance to work in the field of educational psychology.

AHLSTRÖM, K.-G. (1961). *Transfer and Communality: An Experimental Study of Learning*. Studia Scientiae Paedagogicae Upsaliensai, Scandinavian University Books. (Trans. Read, A.). No price given.

This work was intended to test two hypotheses, both of which were concerned with the extent to which ease in learning by rote is dependent upon past experience with similar or identical material. Transfer effects were tested in two ways: (i) by observing the effect of this communality on the first attempt at learning the new material and (ii) by measuring how far the three parameters of the learning equation were changed by the amount of communaility (or similarity) in the two sets of material.

As was predicted, the parameters for the asymptote and rate of learning were invariant with the communality, but the number of correct responses at the first (non-reinforced) trial in the new material varied positively with the amount of communality.

The author considers that inconsistency of results, in experiments designed to study transfer of training, can arise from the methods by which similarity is determined. Accordingly, he systematically established degrees of communality in his response material by the use of transformation matrices in conjunction with tables of random numbers. By this method the proportion of overlap is indicated by the product of the figures in the main diagonal, thus enabling the measure of communality to be based on objective probability values. Material for twelve experimental groups was derived in this way with the proportion of communality ranging from 1·0 to 0·001. His undergraduate subjects were randomly distributed to the twelve groups, and their equivalence for retention tested by analysis of variance, after learning the first set of material.

For those engaged in research in this field, the book will be of interest mainly as a study in methodology, and in particular, of the snags which arise when the subject's strategy of learning is not controlled M. E. HIGHFIELD.

BARNES, KENNETH C. (1962). *He and She*. Penguin Books, pp. IX+203, 3s. 6d.

This is a book about sexual behaviour written for young people in language which they can understand. Although it aims to instruct as well as explain, it is refreshingly free from dogma of any kind, and offers a personal and humane interpretation of moral and religious problems. A teacher who wishes to find material which is suitable to use in discussions with older pupils could not find better help or guidance This is not surprising, as Mr. Barnes is the headmaster of a co-educational school who writes as a teacher from his experiences in talking to boys and girls The book is, in fact, written mainly for boys with the aim of helping them to understand the changes which occur in both sexes at puberty and adolescence in relation to the social patterns of permitted sex behaviour. It should, however, reach a much wider audience and contains much which would help parents as well as children.

Above all, it is a book which could be most valuable for students in teacher-training courses. In the writer's experience, seminars on sex education are attended mainly by those who have already achieved some degree of emotional security through marriage or engagement. The use of this book as a basis for seminar work would help all students who are themselves groping towards sexual maturity. In addition, it is an outstanding example of good communication which could well be copied by psychologists. P. K. POPPLETON.

GARTON, MALINDA DEAN (1961). *Teaching the Educable Mentally Retarded: Practical Methods.* Springfield, Illinois: Charles C. Thomas, xiv+223, £3.

This is avowedly a non-technical book which concentrates on detailing practical steps in teaching various subjects to educable retarded children. The aims of this form of teaching and the limitations of these pupils are only briefly sketched in. It is true that these matters are discussed very fully in other publications, but it is questionable whether practical suggestions without a fuller expression of principles is adequate. For example, a chapter on the use of the sociogram is sandwiched between one on the bulletin board and one on arithmetic. It seems both odd and misleading to stress this one technique without referring to all the other methods for observing and understanding pupils, and without explaining its function as part of a comprehensive assessment of the child. Moreover, it is difficult to appreciate why some topics are discussed in detail, while others are only itemised. For example, activities important in the preparation for school leaving are listed under the headings of promotion of social competence, economic efficiency and self-realization, but the relative importance of particular activities and the methods of teaching them are not discussed. In contrast, arithmetic is dealt with very thoroughly, taking 64 pages which is twice the size of the next largest chapter. Reading is considered in about seven pages, which is less than the amount given to handwriting.

In brief, although there are some useful practical suggestions in the book, it does not as a whole give a sufficiently broad and comprehensive account of the aims and methods of educating mentally handicapped children. R. GULLIFORD.

HALL, OLIVE A., and PAOLUCCI, BEATRICE (1961). *Teaching Home Economics.* London and New York: John Wiley and Sons, pp. xiii+397, 53s.

This book is intended for teachers in training and practising teachers.

Although divided into four sections, it falls into two parts. The first part, covering the nature of Home Economics and the Needs of Individuals and Families, deals with scope and objectives. It leaves no doubt that the course is primarily concerned with homes; not the Home in any abstract or ideal sense but the homes from which the teacher's pupils come, the homes she is expected to contact and visit, the families she will know and work with. The teacher's job starts and finishes there though she has her classroom elsewhere. Her task is to change family behaviour.

Getting to know the pupils and their families, recording and tabulating information about them appears to involve more work than British teachers would have time for and more co-operation from parents than they would expect to find.

The content of the Home Economics course embraces Family and Child Development, Home Management and Family Economics, Food and Nutrition, Clothing, Textiles, and related Art, Housing, Home Furnishing, and Household Equipment. The programme is stimulating and imaginative but students in training and inexperienced teachers would need more practical guidance than is given in implementing it.

The second part of the book is concerned with learning and the appraisal of learning and is largely an enumeration of teaching and evaluation techniques. There is a conscientious but perhaps misguided attempt to acquaint the student with all available classroom procedures and teaching aids: so much advice could stifle spontaneity.

The attempt to exploit current psychological knowledge for the improvement of Home Economics teaching is commendable : there is a certain incongruity in the application of academic learning theories to the arts of dating and baby sitting.

The style will worry some English readers. There are lengthy definitions of simple terms, interminable lists of recommendations and queries, and a multiplicity and complexity of assessment techniques likely to daunt an experienced teacher.

Yet this book has something to offer teachers of Home Economics in this country. Its programme of studies is sound : it relegates craft skills to a subordinate role in home making and emphasises family relationships, preparation for marriage and parenthood.

I would not recommend it for beginners. It would rather be of interest to practitioners who want a new angle on their own subject or another look at teaching methods, but have enough experience to be discriminating in following its recommendations. It should have a place in a college library. OLIVE METCALFE.

HERBERT, W. L., and JARVIS, F. V. (1961). *Dealing with Delinquents*. London : Methuen, pp. 208.

GIBBENS, T. C. N. (1961). *Trends in Juvenile Delinquency*. Geneva : W.H.O. Public Health Paper, No. 5, pp. 56.

Both these works make it possible for students of forensic psychology to familiarise themselves quickly with the background and current trends in criminology. The W.H.O. pamphlet by Dr. Gibbens is written in his usual clear style, and gives briefly facts and recent figures, for instance on such topics as sexual and violence offences. He points out that whereas sexual offences by juveniles do not appear to have appreciably increased in many countries, they have done so in England since the last war. In contrast, offences of violence have increased world wide in recent years, as has female juvenile crime (though it is still proportionally less than male criminality). In his chapter on Present Trends in Prevention, the author draws together current information on prediction studies and on community, medical educational and other projects. He recalls the sad truth that the first casualty of war, and the one which takes longest to recover is education, especially in badly war-torn countries where delinquency rates are highest ; these decrease proportionally with the re-establishment of educational services. In the chapter on Treatment and After-Care, he lists a series of recent experiments.

Dealing with Delinquents, though unpretentious, is nevertheless, a useful book written by two obviously very experienced Probation Officers. They display their writing skills by reducing the vast current literature on delinquency to easily absorbable reading, especially for relative newcomers to the field. Their bibliography alone shows that not only have they selected the major works of the past (e g., on aetiology of crime), but also current books on the subject are given places of honour. They aim at supplying a practical and simple text, and have succeeded in their modest way of doing so. For instance, their two chapters, especially relevant to this journal, are instructive (" The Young Delinquent in School and Community " and " The Teacher and the Delinquent "). They give hints for teachers on how to spot pre-delinquency in time, based not only on Dr. Stott's special tests, but also on general observation of such well-known indicators as truanting, anti-social behaviour, educational retardation, etc. Generally, their approach to the problem is that of the skilled P.S.W.—Probation Officer, who from his knowledge of Freud, is aware of his own psychodynamics as well as those of his clients. The authors go beyond this, however, by demonstrating that such and similar knowledge becomes highly specialised when applied specifically to delinquency as distinct from applying it to such other deviant behaviour as neuroticism. R. G. ANDRY.

ISAACS, N. (1961). *The Growth of Understanding in the Young Child.* London: E.S.A., pp. 41, 3s. 6d.

This brochure, sub-titled, " A Brief Introduction to Piaget's Work," competently fulfils its author's intention to provide for working teachers, first, a thumbnail sketch of the whole story as Piaget presents it, and subsequently " to dwell more fully on the happenings of the Infant School phase." In a final section, Mr. Isaacs goes further in presenting his own balanced interpretation of the relevance of Piaget's work to educational practice, and it is this concluding essay which may well arouse the liveliest interest in those readers already familiar with Piaget's own writings.

The initial ' thumbnail sketch ' is as lucid and accurate an exposition as one could expect in the space of some half-dozen pages, but the succeeding central section of the brochure will undoubtedly be the most meaningful to those students approaching Piaget's ideas for the first time. This describes in concrete fashion the actual experimental situations devised by Piaget to investigate the notions of young children about number, distance, length, area, order, speed and time. This section could well lead on to practical work by students with children and thus to greatly increased insight into the nature of children's thinking.

There is an early passing reference to work being done in this country along Piagetian lines but, apart from the inclusion of Professor Peel's " The Pupil's Thinking " in the short bibliography, no indication is given of the extent of this work. Nevertheless, students in education departments will doubtless have their attention drawn, for example, to the relevant articles in recent issues of this *Journal*, and to Dr. Lovell's newly-published *The Growth of Basic Mathematical and Scientific Concepts in Children*, and it is clear that Mr. Isaacs' slender and unpretentious volume will provide a helpful first step for those striving towards an understanding of the fundamental work of the Geneva school. E. FISK.

KERSHAW, JOHN D. (1961). *Handicapped Children.* London: William Heinemann Medical Books, pp. x+228, 21s.

Many specialist books are available on the treatment and education of each of the major groups of handicapped children, but surprisingly, there has been no recent book in this country which attempted to survey the problems and needs of handicapped children as a whole Yet many workers in this field have to deal not with any one group, but with several. Moreover, both the greater incidence of multiple disabilities and the increasing recognition of the fact that many problems are common to all handicaps point to the need for a book which includes within its scope all the categories of handicapped children.

Dr. Kershaw's book meets the need very well. It is readable and non-technical, giving due weight to medical, psychological and educational aspects. The first seven chapters discuss topics which are relevant to all handicaps—public attitudes to the handicapped, the attitudes and problems of parents of handicapped children, the principles of assessment, the aims of special education, the need for recreation and for work. Each of the remaining ten chapters is devoted to a particular handicap. The nature and causes of the disability and the essentials of treatment and education are outlined A summary of the present state of knowledge and opinion is enriched by the observations and comments based on a considerable experience as a school medical officer.

It is likely with such a book that specialists in particular fields will feel that there is much more to be said, but the author has done a good job in outlining the essentials and in presenting them in a way which never loses sight of human factors in the problem. It is a book which can be recommended as desirable reading for all those who have to participate in making decisions about handicapped children. It is also to be strongly recommended as reading for teachers in training. R. GULLIFORD.

LOVELL, K. (1961). *The Growth of Basic Mathematical and Scientific Concepts in Children.* U.L.P., 15s.

At the present time there is a widespread demand from teachers and students for books about the work of Piaget which describe his experiments succinctly and expound his theory of the development of concepts in a straightforward but critical fashion. Dr. Lovell's book will help to satisfy this demand. His chief aim is to aid students and teachers to understand the development of concepts in their pupils. He discusses briefly how concepts are formed and examines critically some apparatus designed to aid children in forming numerical concepts. In line with Piaget he stresses the importance of activity in achievement of concepts; he suggests that kinds of apparatus which appeal chiefly to visual perception of relationships in relatively rigid structures, are less efficient than those which can be combined in various ways to match the flexibility of mental operations. In the greater part of the book many experiments made by Piaget and Inhelder, or similar experiments made by other investigators, are considered. These show the development of concepts of number, weight, time, space, length and measurement, and area and volume. Lovell draws attention to the evidence that concepts are first understood in specific situations and only later become generalised. He suggests that intuition and experience play a larger part in their development than Piaget allows.

The author's other aims are to increase his readers' understanding of some mathematical concepts and to encourage them to reflect on their activities in teaching. To achieve these aims the logical basis of the natural numbers is discussed (Ch. 2) and the extension of the laws of arithmetic from natural numbers to other numbers is described in an Appendix (Ch. 11). Chapter 2 is at an elementary level but, even so, the statement "more important axioms are called theorems" is inaccurate. Definitions of concepts of quantity are taken from the dictionary and some history of their measurement is given. Chapter 11 is said to be included both as a good example of the mathematical process of generalisation, and because this knowledge is now recommended for students by lecturers in training colleges It should be of interest to teachers but its application to teaching may not be immediately apparent.

The index of the book is very brief and the omission of full stops after all abbreviations is a somewhat disconcerting innovation.

Despite these criticisms this is a useful book. It may well encourage teachers to examine further the concepts which they teach and to inquire more knowledgeably into their pupils' understanding of them. RUTH BEARD.

LURIA, A. R. (1961). *The Role of Speech in the Regulation of Normal and Abnormal Behaviour.* London: Pergamon, pp. 110, 50s.

This contains the three Special University Lectures given by Professor Luria of Moscow, at the University of London, in 1958. Special justification is pleaded by the publishers for the high price and relatively low quality of production, but more attention should have been paid to the linkage between illustrations and text which is often obscure.

Professor Luria is well known to psychologists, most of whom are aware of the importance of his work and will welcome this publication in English. It is scarely less important for educationists who will find it difficult reading but will be amply rewarded for their effort. A rigorous experimental approach to the very simple activities of which a young child is capable has yielded findings which greatly extend those of Piaget and could form the basis of an educational method, reminiscent of Montessori but relating motor activity to language.

Successive stages in the acquisition of speech are described. In the early stages, speech cannot influence motor activity which is regulated by feed-back systems. Later, the child's speech initiates activities but this excitatory function is independent of the meaning of the words used. With further language development, speech takes over a regulative function dependent upon meaning and, ultimately, becomes internalized as an essential component of thought. The final lecture is devoted to pathological modifications of this sequence of development. H. GWYNNE JONES.

MEREDITH, P. (1961). *Teach Yourself: Learning, Remembering and Knowing.* London : English Universities Press, pp. 174, 6s.

This recent addition to the " Teach Yourself " Series is both an account of the psychology of study and a practical guide to strategy and tactics for the solitary student. Of the three processes mentioned in the title, the stress is laid on 'knowihg' in the sense of understanding, and much use is made of the concept of the ' schema.' The main theme is best summed up by a quotation :

> " . . . (*a*) book does most of it's work after it is shut It's work is to add to your stock of schemas. The actual reading only sows seeds. The mind itself has the essential horticultural job of tending these seeds until they burst into flower."

Professor Meredith is obviously a master gardener and, although his advice may, at times, appear a little impractical to the part-time student preparing for an examination, it could be of great value to anyone grappling with a new subject, with or without the aid of a tutor. It may be equally valuable to many who teach and write. Humour is not lacking and is particularly evident in a chapter devoted to the battle between reader and author and the traps of ambiguity and obscurity laid by the latter.

H. GWYNNE JONES.

O'CONNOR, N. (edit.) (1962). *Recent Soviet Psychology.* Oxford, London, New York and Paris : Pergamon Press, pp. 334, £3 10s. 0d.

The price which the publishers have seen fit to put on this volume seems excessive and will naturally tend to deter those who might otherwise buy it. This is regrettable as the book deals with subjects of interest both to psychologists and educationalists.

Following a Preface by Dr. R. H Thouless, there is an introduction by the editor, which gives an outline of some of the characteristics of contemporary Soviet psychology. Emphasis isplaced on three important traditions, stemming respectively from Pavlov, Vygotski, and Uznadze. Mention is made of a preference for physiological formulations, a negative attitude towards mental testing, and the relative neglect of social psychology. The main part of the book comprises sixteen chapters written by Soviet psychologists themselves. Among these is Luria, together with several other leading investigators whose names are becoming known in Britain : Teplov, Ye. N. Sokolov, Leontiev, Galperin (all of Moscow), and Ananiev (Leningrad) Some of the contributors, like Luria and Nebylitsyn (Moscow), and Natadze (Tbilisi), have already published in British journals, though all contributions to this volume—selected mainly from articles in *Voprosy Psikhologii* (Problems of Psychology)—are translations from the original Russian or Georgian. The book is divided into three sections · Physiological, Speech and Education. In the physiological section is a contribution by Teplov which gives one leading Soviet psychologist's standpoint on the problem of individual differences, and which reports researches of his own laboratory into this field . studies based on Teplov's modifications of Pavlov's original typology.

All three sections are likely to be of interest to readers of this *Journal.* A developmental orientation is apparent in the treatment of such subjects as tactile perception, speech, language, and thinking. When the Head of the Moscow University Department of Psychology—A. N. Leontiev—writes about " Learning," his emphasis is both educational and developmental. The chapter by Luria reports some highly original experimental studies of the development of voluntary movements in young children, and another leading investigator Galperin—jointly with N. F. Talyzina—reports on some interesting research into the teaching of elementary geometrical concepts. Several of the chapters reflect the concern of Soviet educational psychology with the learning problems of the child within the school. Very little is available in English translation of the work of the Georgian school, stemming from the earlier researches of Uznadze. Two chapters of this book are of special interest, for this reason In one of them I. T. Bzhalava deals with Uznadze's central concept of ' set ' which—together with the ideas of Pavlov and Vygotski—has greatly influenced contemporary Soviet psychology The second paper, by R. G. Natadze, is one of the most interesting chapters in the book. It surveys the impressive body of research conducted at Tbilisi under Uznadze's influence, into such subjects as language, the process of naming, concept development, and the development of comprehension in the pre-school child.

This book performs a valuable service in making available reports of recent Soviet work, including some vigorous developments in educational psychology. As Dr. Thouless suggests in his Preface, its differences of emphasis and orientation may have something to contribute to the enrichment of our own work

PETER MCKELLAR.

TANNER, J. M. (1961). *Education and Physical Growth.* University of London Press, pp. 144, 12s. 6d.

I have long regretted the striking lack of communication and liaison between doctors, especially paediatricians, and teachers, between whom there is so much in common Each, I am sure, has a great deal to contribute to the knowledge of the other. This book, by Dr. Tanner, is a valuable contribution to the teaching profession by the leading British expert on the physical growth of children. The book originated in a series of lectures given at the Institute of Education in the University of Leeds.

The author has attempted to pick out the main facts concerning physical growth which have a bearing on educational practice. He lays particular emphasis on the relationship between physical development and intellectual and emotional problems. Amongst the interesting matters discussed, one may mention the following: the emotional and educational implications of the fact that maturation tends to be earlier in girls than in boys: the secular trend toward earlier maturation: the association of advanced physical development with advanced mental ability: the emotional and educational difficulties associated with an unusually early or unusually late onset of puberty: the importance and anatomical basis of maturation of the brain in relation to learning, and to Piaget's stages of mental functioning: the advantage of early physical maturation in boys in relation to dominance over their fellows: the prediction of adult size from the growth in childhood: and the factors affecting physical growth. The author discusses many other fascinating questions concerning physical growth

I cannot help feeling that some of the implications of the different levels of maturation in children of the same age are somewhat more theoretical than practical. Tanner suggests, for instance, that school leaving age should be governed more by the developmental age than by the chronological age; that the age of learning arithmetic and reading should be related to the level of maturation achieved; and that developmental homogeneity should be aimed at in physical education. After discussing the differential rate of maturation, he writes: " It behoves us to fit our educational system, in theory and practice, to these biological facts." Tanner himself realises the difficulties in carrying these aims into effect. One feels, however, that he has not fully considered some of the emotional implications of matching children of the same level of physical maturation. For example, mentally superior children who have moved out of their age group in order to do work more suitable for their mental age, or small boys with average intelligence whose physical growth is much less advanced than others in their form, will be likely to find that they lag considerably behind their form fellows in sports and physical education: but such boys are not likely to be pleased to have to leave their form fellows in order to join boys in a form a year or two behind theirs for physical education and games. As a paediatrician, and not an educationalist, I would have thought that such treatment would increase any feeling of inferiority which they felt, and not reduce it. It is a matter of opinion whether it is more important to match mental age or physical age when the two do not parallel each other Anyway, I doubt whether an average child suffers any significant emotional trauma by being smaller than others in his form, or less good than others at gym, especially if he is better than they are in other subjects.

Dr. Tanner's book is a thoroughly scientific, readable, interesting and important book which all teachers and others concerned with children should read and buy for their bookshelf. It has the additional merit of being inexpensive.

R. S. ILLINGWORTH.

TERMAN, L. M., and MERRILL, M. A. (1960). *Stanford-Binet Intelligence Scale. Third Revision, Form LM*. Harrap, Manual, pp. 406, 21s. Card material, 21s., Case of Test Material, 160s. + 32s. purchase tax.

An anglicised version of the 1960 Terman and Merrill Test is now available in this country. It is claimed that the best sub-tests from the old L. and M. forms of the 1937 version have been selected and combined into a single scale to be known as the L-M Form. The selection of items was based on the testing of a representative sample of nearly 4,500 American school children and on data presented by a number of research bodies concerned with child development. In principle, the 1960 scale is the same as the previous versions in that it is still an age scale so devised that the average mental age agrees closely with chronological age. The location of sub-tests is on grounds of their difficulty defined in terms of the percentage of children passing. Within this structure, however, the new norms depart from older practice in that they are no longer strictly I Qs. but standard scores having a mean of 100 and standard deviation of 16. There are tables which convert the mental age score into the standard score direct.

There is no doubt that the material in the new form is more appropriate to present-day conditions. This writer found it pleasant to use and easy to pick up from an intimate knowledge of the 1937 version. As yet, of course, we know nothing of the relevance of the norms for British children or whether corrections like that of Roberts and Mellone will be needed. A report by the Division of Professional Psychologists on field trials of the test will be welcomed L. B. BIRCH.

THIBAUT, J. W., and KELLEY, H. H. (1959). *The Social Psychology of Groups*. New York: Wiley and Sons; London: Chapman and Hall, Ltd., pp. 313, 56s.

This book is the result of a period spent by the authors at the Centre for Advanced Study in the Behavioural Sciences, Stanford, California, which enabled them to " take a long and fresh look at the basic theoretical problems of the small group field." It offers a new approach to the problems of interdependence, attraction to the group, power and control, status evaluations, social norms, etc., including an attempt to give delimited conceptual definitions to these and other terms commonly used in the study of groups. The new proposals in part arise from conceptions already accepted in such areas as economics, learning theory and sociology.

For instance, Von Neumann and Morgenstern, in their *Theory of Games and Economic Behaviour*, Princeton, Princeton University Press, 1947, assume " the consumer desires to obtain a maximum of *utility* or satisfaction and the entrepreneur a maximum of profits," which is closely paralleled by the basic concept of Thibaut and Kelley that the interactions and their consequences in a group of individuals can ultimately be expressed in terms of a *reward-cost* matrix; in common with the economists, also, the latter begin by considering the dyad, or two-person relationship, and recognise the increasing complications of larger groups due to the presence of sub-groups and coalitions. Nevertheless, they have shown that the analysis of the dyadic relationship can do much to illuminate the problems of the larger group, and use them to examine the concepts of ' role,' ' norms,' ' power,' ' group cohesiveness,' and ' status,' with greater precision than much of the former work on these subjects.

Perhaps the weakness of the book, freely admitted by the authors, is its inability to produce specific experimental evidence to test the hypotheses formulated, though it is unusually rich in suggesting areas for future research. It illustrates freely from existing research literature in the field, and admits that at times it has resorted to anecdotes and authoritative statements in lieu of better evidence. Nevertheless, it makes a genuine effort to analyse the interrelatedness of diverse research findings, the summaries of which, alone, are very useful in collected form.

The book is very well produced, with good clear diagrams of reward-cost matrices, etc.; the style is lucid and scholarly throughout, and the presentation greatly aided by the presence of summaries at the ends of the chapters.

D. M. EDWARDS PENFOLD.

TIZARD, J., and GRAD, T. C. (1961). *The Mentally Handicapped and their Families.* London: Oxford University Press, pp. x+146, 28s.

This monograph by a psychologist and a psychiatric social worker, describes the social and psychological problems encountered by the relatives of severely sub-normal patients In a social survey dealing with 150 London families having their child at home and 100 families with a child in an institution, the effects upon the families are investigated The data were obtained by interviews and from records and use was made of the Vineland Social Maturity Scale. Many of their conclusions support now by facts and figures assumptions made in the past, but there are also other findings which destroy some prejudices and widely held beliefs. For example, the authors came to the conclusion " that the middle class are most likely to request institutional care " rather contrary to the view that it is the mental defective child of the poor who most readily finds his way into an institution. The data also failed to support the view that " large families are more likely to send their mentally defective child to institutions " and suggest " that in general it was in working class households and in large families that the best adjustments to the child who is handicapped " were found.

The most important chapter is, however, the section which makes certain generalisations regarding the medical, social, educational and employment services for the mentally subnormal. The recommendations made are not revolutionary or novel; they have been voiced by the progressive workers in the field for many years—with little general effect. Yet, it is helpful to have available now a factual account giving an approximate idea of the size of the various problems involved. Some points in the survey help, however, to emphasise certain aspects of remedial services which have not led to adequate action. First, " the majority of defectives do not impose on their families intolerable problems of management," secondly, " the great majority of parents wish to keep their mentally subnormal children at home " and thirdly, " placement in hospital is best postponed." This has important repercussions on the design of the now developing community services which should be organised to *help the parent in the Management of the defective at home,* rather than simply provide places for occupation during the day The present monograph brings out the real needs of the parents for more direct sympathetic help with the social, emotional, educational and vocational problems in the day-to-day management of their children.

There are probably not more than five studies available on the problems of management at home of the severely retarded The study helps, therefore, to make a very noticeable gap in our knowledge less glaring and supplies at least some of the important data on emotional and social aspects which are required when planning remedial services for the subnormal. In these times of change in the subnormality services this book comes most opportunely and must be regarded as essential reading for all those who work with the mentally subnormal. H. C. GUNZBURG.

TRASLER, GORDON (1962). *The Explanation of Criminality.* London: Routledge and Kegan Paul, pp. 134, 20s. net.

Of the making of books there is no end, and this saying seems particularly appropriate when we consider crime and criminality. Unfortunately, quality seems inversely correlated with quantity. Trasler's book is a most welcome exception to this generalization; in all the welter of psychoanalytic, sociological, psychiatric, statistical, and 'judicial' tomes written on this subject, this modest, short and well-written essay stands out as exemplifying a truly rational, scientific approach to the problem. Trasler starts out by assuming that while heredity and constitutional factors no doubt play some part in the predisposition of some individual towards crime, yet the major factor must be found in social learning and conditioning. He goes on to argue that the contributions of learning theory to this field could be very great, and suggests ways of explaining criminal behaviour in these terms. He bases himself essentially on Mowrer's model of the socialization process which is well and concisely explained with reference to some experimental studies.

Trasler then goes on to consider individual differences, particularly those related to a person's conditionability, and from a consideration of these various factors he arrives at a theoretical scheme which is set out in four formal propositions (p. 63). Trasler is aware, and emphasizes particularly, that these individual differences are always manifested in a social context, and that the final outcome will be determined by both acting together. Thus, it is important not only that an individual should condition easily, but also just what the conditioned and unconditioned stimuli are; this obvious point has often been neglected by critics of this approach. From his formal propositions, Trasler then makes a number of deductions, and produces some evidence to show that these are indeed valid. Other deductions are given in the form of concepts, and no validation studies are available. In one or two instances it is possible to find fault with his argument. Thus, he states (p. 74) that "social conditioning will be most effective where sanctions are applied consistently and reliably." This appears to be a plea for 100 per cent. reinforcement, yet there is good evidence to show that extinction is quicker with 100 per cent reinforcement than with partial reinforcement. There may be special reasons for not applying this general rule in this case, and it might have taken Trasler too far to discuss this in detail; nevertheless, the point might be worth considering. Altogether the book, like all good books, raises more problems than it solves, but at least they are the right problems, rather than the traditional pseudo-problems, and Trasler is to be congratulated on a contribution which is breaking new ground and setting a new and high standard for those who wish to apply psychology to the criminal field.

H. J. EYSENCK.

PUBLICATIONS RECEIVED

Mention of a book in this list neither implies, nor precludes, a later review.

J. Church (1961). *Language and the Discovery of Reality.* New York: Random House, pp. xv+245, no price.

S. Wiseman, *et al.* (1961). *Examinations and English Education.* Manchester University Press, pp. xx+188, 21s.

J. M. Tanner (1961). *Education and Physical Growth.* University of London Press, Ltd., pp. 144, 12s. 6d.

P. T. Young (1961). *Motivation and Emotion.* John Wiley and Sons, Inc., pp. xiv+648, no price.

W. H. Holtzmann, J. S. Thorpe, J. D. Swartz, E. W. Herron (1961). *Inkblot Perception and Personality.* University of Texas Press, U.S.A., pp. vii+417, $8·00.

R. L. Thorndike and E. Hagen (1961). *Measurement and Evaluation in Psychology and Education.* J. Wiley and Sons, Inc., 2nd edition, pp. 602, 58s.

T. R. Fyvel (1961). *The Insecure Offenders.* Chatto and Windus, pp. 347, 25s.

A. N. Frandsen (1961). *Educational Psychology.* McGraw-Hill Pub. Co., Ltd., pp. xii+610, 58s.

M. Langdon (1961). *Let the Children Write.* Longmans, pp. viii+72, 6s. 6d.

A. J. Ungersma (1961). *The Search for Meaning.* G. Allen and Unwin, Ltd., pp. 188, 25s.

H. Guntrip (1961). *Personality Structure and Human Interaction.* Hogarth Press, Ltd., pp. 456, 45s.

R. P. Lynton (1961). *The Tide of Learning.* Routledge and Kegan Paul, pp x+228, 25s.

H. Solomon (1961). *Studies in Item Analysis and Prediction.* Oxford University Press, pp. ix+310, 70s.

World Health Organisation (1961). *Teaching of Psychiatry and Mental Health.* W.H.O. Public Health Papers No. 9, pp. 186, 10s.

American Foundation for the Blind (1961). *Adjustment to Visual Disability in Adolescence* American Foundation for the Blind, New York, pp. 259 $4.50.

J. E. Sadler and A. N. Gillett (1961) *Training for Teaching.* G. Allen and Unwin Ltd., pp. 272, 21s.

L. Blomfield and C. Barnhart (1961). *Let's Read.* Wayne, U.S.A., pp. 470, $7.50.

J. McV. Hunt (1961). *Intelligence and Experience.* New York: Ronald Press, pp. ix+416, $8.0

D Stafford-Clark (1960). *The Psychology of Persecution and Pressure.* The Council of Christians and Jews, pp. 27, 2s.

H. C. Dent (1961). *The Educational System of England and Wales.* University of London Press, pp. 224, 15s.

T. H. Pear (1961). *The Moulding of Modern Man.* George Allen and Unwin, pp. 220, 21s.

Harry Benjamin (1961). *Basic Self-Knowledge.* Health for All Publishing Co., pp. 168, 15s.

Felix Scherke (1961). *Der Politische Charakter.* Stuttgart: S. Hirzel Verlag, pp. 115, DM. 12.60.

Allen D. Calvin, *et al.* (1961) *Psychology.* Allyn and Bacon, Inc., Boston, pp. xix+571, 72s.

K. Lovell (1961). *The Growth of Basic Mathematical and Scientific Concepts in Children.* University of London Press, pp. 154, 15s.

Oscar Krison Buros (edit.) (1961). *Tests in Print.* New Jersey: The Gryphon Press, p.p. xxix+479, $7.0

Arnold H. Buss (1961). *The Psychology of Aggression.* London and New York : John Wiley and Sons, pp. x+307, 60s.

Olive A. Hall and Beatrice Paolucci (1961). *Teaching Home Economics.* London and New York : John Wiley and Sons, pp. xiii+397, 53s.

Nathan Isaacs (1961). *The Growth of Understanding in the Young Child :* London : Educational Supply Association, Ltd., pp. 41, 3s. 6d.

E. Anstey (1961). *Staff Reporting and Staff Development.* London : George Allen and Unwin, Ltd., pp. 96, 12s. 6d.

Georg Schmidt, *et al.* (1961). *Though this be Madness* London : Thames and Hudson, pp. 114, 42s.

Frederick C. Gruber (edit.) (1961). *Anthropology and Education.* Pennsylvania University Press, Oxford University Press, pp. 123, 22s.

Kenneth C. Barnes (1962). *He and She.* Penguin Books, pp 206, 3s. 6d.

Gordon Trasler (1962). *The Explanation of Criminality.* London : Routledge and Kegan Paul, pp. ix+134, 20s.

Morris Ginsberg (1962). *On the Diversity of Morals.* William Heinemann, Ltd., pp. xiv+329, 12s. 6d.

Lilian Hollamby (1962) *Young Children Living and Learning* : Longmans, Green and Co., Ltd., pp. 198, 12s. 6d.

Brian Jackson and Dennis Marsden (1962). *Education and the Working Class.* London : Routledge and Kegan Paul, pp. ix+268, 28s.

Sir Cyril Burt (1962). *Mental and Scholastic Tests* (*4th Edition*). London : Staples Press, pp. xxx+551, 65s.

Oskar Spiel (1962). *Discipline Without Punishment.* Faber and Faber, pp. 171, 21s.

Prudence Hamilton (1962). *The Greatest Role in the World.* New York : Vantage Press, Inc., pp. 107, $2.95.

Unesco (1962). *Reports and Papers in the Social Sciences, Nos.* 14 *and* 15. 7s. 6d. each.

World Health Organisation (1962). *Who Chronicle, Vol. 16, No. 1.* Geneva : World Health Organisation, 1s. 9d.

A. M. Kean, *et al.* (1961). *Researches and Studies No. 22, Dec.,* 1961. University of Leeds Institute of Education, 5s.

Kenneth E. Clark (1961). *Vocational Interests of Non-Professional Men.* Minnesota University Press, Oxford University Press, pp. xi+129, 30s.

C. H. Patterson (1962). *Counseling and Guidance in Schools.* New York : Harper and Brothers, pp. xiii+382, no price stated.

Sir James Gray, *et. al* (1961). *The British Association/Granada Guildhall Lectures, 1961.* University of London Press, pp. 80, 4s. 6d.

A. H Halsey (edit.), *et al.* (1961). *Education, Economy and Society.* New York : The Macmillan Co., pp. ix+625, 56s.

J. Maxwell (1961). *The Level and Trend of National Intelligence.* University of London Press, pp. 75, 5s. paper, 7s. 6d. boards.

N. O'Connor (edit.) (1961). *Recent Soviet Psychology.* Oxford : Pergamon Press, Ltd., pp. 334, 70s.

Thomas E. Jordan (1961). *The Mentally Retarded.* London : Prentice-Hall International Inc., pp. x+355, 50s.

Hyman S. Lippman (1956 and 1962). *Treatment of the Child in Emotional Conflict.* London : McGraw Hill Publishing Co., Ltd., 2nd edition, pp. xii+367, 74s.

Ernest H. Hutten (1962). *The Origins of Science.* George Allen and Unwin, Ltd., pp. 241, 28s.

Robert Elliot Fitch (1962). *Odyssey of the Self-Centred Self.* George Allen and Unwin, Ltd., pp. 184, 16s.

Robert W. Shields (1962). *A Cure of Delinquents.* London : Heinemann, pp. 191, 21s.

K. M. Evans (1962). *Sociometry and Education.* London : Routledge and Kegan Paul, pp. vii+149, 18s.

R. P. Farnsworth (edit.) (1962). *Annual Review of Psychology—Vol.* 13, *1962* California : Annual Reviews, Inc., pp. ix+602, $7.0

G. Thines (1962). *Contribution a la Théorie de la Causalité Perceptive.* Louvain : Editions Nauwelaerts, pp. 124, FB 140 (2.80 dollars).

Joseph Nuttin (1961). *Psychology in Belgium.* Louvain : Éditions Nauwelaerts, pp. 80, FB 70 (1.60 dollars).

Marie-Therese Knapen (1962). *L'Enfant Mukongo* Louvain : Éditions Nauwelaerts, pp. 204, FB 170.

Proceedings of First International Congress on Ergonomics, Stockholm (1961). *Ergonomics, Vol. 5, No. 1, Jan 1962.* London : Taylor and Francis, Ltd., pp. 339, £2 10s. 0d.

K. R. Eissler (1961). *Leonardo da Vinci.* London : The Hogarth Press, Ltd., pp. xx+375, 65s.

Carl Rogers and G Marian Kinget (1962) *Psychotherapie et Realtions Humaines.* Louvain : Editions Nauwelaerts, pp. 320, FB 240 (5 dollars).

Society for Research in Child Development (1962). *Child Development, March, 1962, Vol. 33, No. 1.* Society for Research in Child Development, Inc., pp. 255, $3.50.

James Maxwell (1961). *The Level and Trend of National Intelligence.* London : University of London Press, Ltd , pp. 77, 7s 6d.

Robert L. Sutherland (edit.), *et al.* (1962) *Personality. Factors on the College Campus.* Texas : The Hogg Foundation for Mental Health, pp. 242, $2.25

Mary Stewart (1962). *The Success of the First Born Child.* The Workers' Educational Association, pp. 19, 2s. 6d.

Lewis M. Terman and Maud A. Merrill (1961). *Stanford-Binet Intelligence Scale—Manual for the Third Revision Form L-M.* London : Harrap and Co., Ltd., pp. 405, 21s.

J. C. Johnson (1962). *Educating Hearing-Impaired Children in Ordinary Schools.* Manchester University Press, pp. ix+120, 25s.

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Statistical Psychology.
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free). Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4., or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s. : $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1., from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to ·
Dr. BORIS SEMEONOFF, Department of Psychology, The University, Edinburgh.

For the *British Journal of Medical Psychology*, should be sent to Dr. JOSEPH SANDLER, 96, Portland Place, London, W.1.

For the *British Journal of Educational Psychology* should be sent to Mr. L. B. BIRCH, Institute of Education, The University, Sheffield, 10.

For the *British Journal of Statistical Psychology*, should be sent to Mr. JOHN WHITFIELD, Department of Psychology, University College, London, Gower Street, W.C.1.

For the *British Journal of Social and Clinical Psychology* should be sent to Mr. MICHAEL ARGYLE (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford ;

or to ·

Dr. JACK TIZARD (Clinical Psychology Editor), Maudsley Hospital, Denmark Hill, London S.E.5.

For the *Monograph Supplements*, to Professor L. S. HARENSHAW, Department of Psychology, The University, 7, Abercromby Square, Liverpool, 7, whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May and distributed free to members of the Association. To non-members it costs, 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

CONTENTS

PAGE

L. B. Ames : Dr. Arnold Gesell, 1880-1961 101

D. Case and J. M. Collinson : The Development of Formal Thinking in Verbal Comprehension 103

J. McLeod : The Estimation of Readability of Books of Low Difficulty. 112

S. Jones : The Wechsler Intelligence Scale for Children Applied to a Sample of London Primary School Children . 119

R. R. Skemp : The Need for a Schematic Learning Theory . . 133

M. D. Vernon : Specific Dyslexia 143

H. B. Young, R. Tagiuri, G. Tesi and G. Montemagni : Influence of Town and Country upon Children's Intelligence 151

A. Royds : Intelligibility of Schools Television Programmes . 159

J. A. Downing : The Relationship Between Reading Attainment and the Inconsistency of English Spelling at the Infants' School Stage 166

Symposium : Educational Research To-day :

E. A. Peel : I.—Priorities in Educational Research . . . 178

G. E. R. Burroughs : II.—Co-operative Research in Institutes of Education 182

J. B. Biggs : The Relation of Neuroticism and Extraversion to Intelligence and Educational Attainment . . 188

R. Lynn : Comments on the Article by J. B. Biggs . . 196

Summaries of Researches Reported in Degree Theses. 200

Book Reviews 207

Publications Received 216

28 JUL 1962

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1962

VOL. XXXII, PART 3

Price Fifteen Shillings

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by

L. B. BIRCH

Associate Editors:

P. E. VERNON and DORIS M. LEE

Advised by

SIR CYRIL BURT

R. A. C. OLIVER

E. A. PEEL

SIR FRED J. SCHONELL

C. W. VALENTINE

S. WISEMAN

Vol. XXXII, Part III

The British Journal of Educational Psychology is issued by the British Psychological Society and the Association of Teachers in Colleges and Departments of Education, once a term, in February, June and November. The subscription is 30s. a year post free payable in advance, or 15s. each number.

Subscriptions and Orders should be sent to "The Manager, B.J.E.P. Department, Methuen and Co., Ltd., 36, Essex Street, Strand, London, W.C.2." Members of the British Psychological Society and members of the A.T.C.D.E. receive the *Journal* on special terms through the Societies.

Papers for publication should be sent to L. B. Birch, The Institute of Education, The University, Sheffield, 10. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to L. B. Birch, at the above address. *Advertisements* and *Subscriptions* for this *Journal* should be sent to Manager, B.J.E.P. Department, Methuen and Co., 36, Essex Street, Strand, London, W.C.2.

Published by

METHUEN AND CO., LTD., 36, ESSEX STREET, STRAND, LONDON, W.C.2.,
for the British Journal of Educational Psychology, Ltd.

Registered Offices: 42, Hill Street, Birmingham, 5. This address should *not* be used for letters or books intended for the Editor, Authors, or the Publishers.

On behalf of

The British Psychological Society and the Association of Teachers in Colleges and Departments of Education.

All rights reserved.

PRINTED BY THE BIRMINGHAM PRINTERS, LTD., 42, HILL STREET, BIRMINGHAM, 5.

INCREASED SUBSCRIPTIONS

The Directors of the *Journal* very much regret that rising costs of production and indications of further increases in the near future, make it imperative that the price of the *Journal* shall be raised. The subscription rate for the new Volume starting with the February number, 1963, will be £2 a year.

For some time, financial stringency has made the Editor's job specially difficult as a larger than usual number of diagrams and tables or a few extra pages could bring financial loss. It is hoped now that it will be possible to expand the *Journal* somewhat, to include more reviews and abstracts of theses and to be more tolerant to authors who need numbers of complicated tables.

L.B.B.

A FURTHER APPROACH TO ATTITUDE SCALING

By R. S. ADAMS.
(*Department of Education, University of Otago, Dunedin, New Zealand*)

SUMMARY. In an attempt to find a simple, uncomplicated and effective way of measuring attitudes that might prove useful to the classroom teacher, a series of experiments was carried out at a teachers' training college. A self-rating scale which required the conceptualising of attitudes in numerical rather than verbal terms was devised as the principal instrument and the results of its application were compared with those from two traditional attitude scales constructed for the same purpose. The self-rating scale proved as effective as the other scales and in addition, gave evidence of having validity against a behavioural criterion.

I.—INTRODUCTION.

To the average classroom teacher the measurement of attitudes has been only of passing interest. There are two principal reasons for this. The first relates to the nature of the particular measuring instruments available, the second concerns the relevance of attitude test results to the practical teaching situation. In the first place, it has been pointed out that projective tests require specialised administration skills ; that self-rating scales are of doubtful reliability ; and that the traditional Thurstone and Chave (1929) and Likert (1932) attitude scales are time consuming to construct, cumbersome to score and somewhat specific in their applicability. In the second place, the kinds of tests that, on face value appear most suitable for classroom use have had their validity called into question, principally on the grounds of " lack of correspondence between what the subject describes as his attitude and the actual overt activities in which he participates " (Klineberg, 1954). There is then, an hiatus between attitude-measured-by-a-test and attitude-in-action. While such an hiatus needs be no deterrent to the researcher, to the classroom teacher, preoccupied as he is with the manifestations of children's behaviour, its existence is a distinct disadvantage. Yet, the development of ' right attitudes ' in children is a teaching aim that is accorded universal recognition. Consequently, the devising of a manageable test which relates more to attitude-in-action would seem a worthwhile undertaking.

II.—CONSTRUCTION OF THE SCALE.

The experiment reported below was concerned with developing a simple but practical scale to assess the attitudes of teachers' college students towards their courses of study. Basically, the scale was a self-rating scale. However, there is one feature in particular that distinguishes it from the more usual of the genus. The students were required to respond with a number of numerical ratings. In other words they were constrained to conceptualise their attitudes not in words, as is customary, but in numbers. The justification for this approach will be considered later, at this stage the instrument itself needs to be described. It took the form of an individual paper and pencil test with the following format and directions.

Self Rating Attitude Scale.

This first test is to measure your attitude towards Physical Education.[*]

You are asked to give yourself a rating between 0 and 100 based on your liking for the subject.

To help you to do this there is below, a graded scale which runs from 0 to 100 mainly in jumps of 10. This is the ' Tens Scale.' Please put a tick somewhere on this scale to indicate the extent of your enthusiasm for Physical Education. If you *like* the subject, your tick should go in the top half of the scale (51–100); if you *dislike* it, it should go in the bottom half of the scale (0–49). The more you like it, then the higher up the scale your tick should go; the more you dislike it, the lower down it should go. If you neither like nor dislike it, you should put your tick in the neutral (50) box. Please do this now.

TENS SCALE

LIKE	☐	91–100
	☐	81–90
	☐	71–80
	☐	61–70
	☐	51–60
NEUTRAL	☐	50
DISLIKE	☐	40–49
	☐	30–39
	☐	20–29
	☐	10–19
	☐	0– 9

There is another scale below, the ' Units Scale.' The numbers in it represent the units from the ' tens box ' in which you have just put your tick. Would you please put another tick in *this* scale. This is to show whether your liking falls nearer to the *top* or the *bottom* end of the ' tens box.'

UNITS SCALE

0	1	2	3	4	5	6	7	8	9

For the remaining tests you are asked to indicate your liking for a number of other subjects.[*] Would you please complete these making sure each time that you follow the same procedure carefully; put a tick in the ' tens box ' first and then in the ' units box.'

On the final page of the test were further directions that enabled the students to score each test, record the results and convert these into an attitude profile. They were also required to indicate which of these courses had been taken voluntarily to a more advanced level as part of their elective programme.

---

[*] The other subjects also tested were, English, Music, History, Geography, Science, Art and Craft, Mathematics and Education. Each was accompanied by an identical diagram.

III.—VALIDITY.

The general format of the test was decided upon after a series of prototypes had been tried out with a small group of advanced education students at the University of Otago. To accomplish the next task, checking the test's validity, two procedures were employed. The first followed the traditional practice of comparing the results from an application of the scale with the results of other tests purporting to measure the same thing. The second required the examination of the test results in conjunction with a behavioural criterion. This latter was supplied from an analysis of the subjects selected by the students for the optional part of their courses. The details of both procedures appear below.

(i) *Inter-test validity.*

Because the form of the self-rating scale was being evaluated rather than the content, only one course of study, physical education, was chosen for inter-test comparisons. As a first step, the making of a further scale to measure attitude towards this subject was undertaken in accordance with the Thurstone and Chave method. The number of statements from which the initial selection was made numbered 150. The judges comprised fifty psychology and education students at the University of Otago. Sufficient items resulted to allow for the construction of two scales with identical mean scale values and similar item, scale and Q values. The two alternative forms (A and B) had twenty items in each.

In addition, a Likert scale was constructed. The statements used were taken from the Thurstone and Chave scales. Eight of the items (the middle ground ones) were rejected because they failed to meet the requirements when the customary ' discriminating power ' criterion was applied. There were thirty-two items in the final scale. The use of identical statements in the Thurstone and Chave test and the Likert tests meant that it could be assumed with more justification that the two tests were measuring the same thing and that differences in result were more likely to be attributable to method than to content.

These two tests were then given to 245 students in their first year at teachers' college. The scores obtained were correlated with each other and also with scores from the self-rating scale given to the same group. Table 1 contains the resultant coefficients.

TABLE 1

INTER-TEST CORRELATIONS.

Tests	r	
Thurstone A—Likert	·77	
Thurstone B—Likert	·80	
S.R.S.*—Thurstone A	·61	
S.R.S.—Thurstone B	·69	
S.R.S.—Likert	·76	N=245

* Self-Rating Scale.

These results do not on face value show the self-rating scale in quite as favourable a light as the others. However, when the mutuality of items in the two verbal scales is taken into account, and as well the ·76 correlation between the self-rating scale and the scale that scored most consistently of all, the Likert scale, it is reasonable to contend that none of the tests appears markedly superior to the others. Furthermore, as instruments in the field of personality measurement, all of the tests fall within the limits of acceptability for inter-test validity.

(ii) *Behavioural Validity.*

The approach adopted here rested on the assumption that students would tend to choose for their optional courses (or ' credits ' as they are designated) the courses of study that they liked best. Thus, if the scale discriminated well, the courses given the highest scores would tend to feature also as ' credit ' courses. It should be noted that ' highest ' scores were identified by the *ranking* each individual student gave the courses. It was *not* assumed (as it is with the Thurstone and Chave and Likert scales) that a particular, arbitrarily selected point on the scale would serve to denote the cutting line between high and low scores. In other words, intensity of attitude here was regarded as peculiar to the individual rather than to the group with which he was fortuitiously associated.

Of the nine courses on which the students returned scores, six were available as ' credits.' The 215 students in this sample had at this stage taken an average of two credits each although nine had taken three and thirty-eight only one. There were twelve incomplete returns and eighteen special case returns that were not applicable. When the responses of each individual student were put in rank order and the positions of the credit subjects identified, the results shown in Table 2 were forthcoming.

TABLE 2

THE DISTRIBUTION OF " CREDIT " SUBJECTS IN A RANK ORDERING OF ALL SUBJECTS.

Ranking	Total number	%	Cumulative %
1st	173	46·6	46·6
2nd	89	24·0	70·6
3rd	53	14·3	84·9
4th	28	7·6	92·5
5th	23	6·2	98·7
6th	5	1·3	100
	371		N=215

The fact that 70·6 per cent. of all choices was given to the two top-ranking subjects indicates that factors other than chance were operating here. Admittedly, in selecting ' credits,' preference for a course would not be the only variable to take into account. Students with a view to their professional competency might deliberately set out to take subjects towards which they were not very favourably disposed. Indeed, the mathematics ' credit ' provides remedial work for those who are particularly weak in school arithmetic and, in fact seven of the twenty-eight choices ranked fifth and sixth were given to mathematics.

Nevertheless, when the general freedom of choice is taken into account, together with the hedonistic inclinations of most students, the results confirmed the earlier hypothesis namely, that if the test did, in fact, measure attitude towards subjects, the students would tend to choose as ' credits ' subjects that ranked high on their test scores.

IV.—RELIABILITY.

As a first check on reliability, some eight months later the initial group used in the inter-test validity experiment was asked to repeat the physical education self-rating scale test. Two factors in particular could have influenced the results of the re-test. First, in the time that had elapsed attitudes could have changed, and second, the scores recorded in the original test might have been remembered. Consequently on the second occasion, after they had completed the re-test the students were asked to indicate whether they thought their attitudes had become ' more favourable,' ' less favourable,' or had undergone ' no change ' in the intervening period. They were also required to record their original score as they remembered it and to indicate how sure they were that they had recalled it correctly by underlining one of the following alternatives: ' quite certain,' ' fairly certain,' ' not too sure,' ' guesswork.'

Of the 190 students who completed this part of the experiment, fifty-four (28 per cent.) indicated a change of attitude—fifty in the ' more favourable ' direction, four in the ' less favourable.' In only five cases was this estimate not confirmed by the direction of change reflected in their test scores. In three of these five cases the scores coincided exactly with the original scores. In the remaining two, the direction was reversed. Test-retest correlations for both the whole sample (190) and the ' no change ' groups (136) were calculated. Table 3 shows the results and in addition shows the reliability coefficients derived from the Thurstone and Chave and Likert scales.

TABLE 3

RELIABILITY OF SCALES.

Scale	Method of Assessment	r	N
S.R.S. (Whole sample) ..	Test-retest	·78	190
S.R.S. (" no change " group)	Test-retest	·83	134
Likert................ ..	Split half	·89*	250
Thurstone and Chave	Comparison of equivalent forms	·71	245

* Corrected for attenuation.

In recording the extent of their confidence in the accuracy of their recall of the original test scores, four students claimed to be ' quite certain,' forty-seven to be ' fairly certain,' seventy-nine to be ' not too sure ' and fifty-nine conceded ' guesswork.' There was one omission. Of the four in the first group, one was actually correct, one missed by a single point, another by two points and the last by nine points. Among those students who expressed themselves as ' quite certain ' or ' fairly certain,' that their attitudes had undergone no change, the average deviation of their second score from their first was 4·7 points. Of the forty cases involved here only four were able to recall their original scores with complete accuracy.

It would seem then that in this particular case the influence of recall of the earlier test results could only be regarded as minimal.

Some further evidence bearing on the reliability of the test is also available. One year later the complete self-rating scale was administered to a different intake of students who were at the same relative stage in their career as were the first group when tested. The students from both groups had been selected for entry to teachers' college according to approximately the same criteria, although it should be noted that the academic qualifications of the second group were on an average slightly higher than that of the first group. The mean scores for all subjects for both groups appear in Table 4.

TABLE 4

Mean Scores of Two Groups.

		Group I—1st Intake (29/11/60)		Group II—2nd Intake (30/11/61)	
		Mean	S.D.	Mean	S.D.
I	Physical Education ..	70·37	15·9	69·88	17·34
II	English	67·55	16·15	68·49	15·32
III	Music	73·60	16 5	72·23	18·49
IV	History	65·70	20·0	65·43	19·46
V	Geography	70·26	16·4	69·64	16·28
VI	Science	60·35	18·3	63·60	19·22
VII	Art and Craft	65·20	18·3	66·57	20·56
VIII	Mathematics ...	59·91	21·15	59·99	23 02
IX	Education	71·30	14·3	74·43	14·08
		N=190		N=179	

Undoubtedly, a number of variables contribute to the formation of attitudes not the least of these involving factors other than general academic attainment, intelligence and personal suitability for training as teachers. It cannot be assumed that two groups selected on the basis of such criteria should necessarily possess identical attitudes. However, there is a prima facie case for predicting that successive groups of teachers' college entrants over at least a short space of time would reflect similar attitude patterns. The data in Table 3 tend to confirm this.

V.—Discussion.

At this stage the evidence available indicates that the kind of self-rating scale devised provides a suitable means for measuring the attitudes of teachers' college students towards their courses of study. Also a comparison of its results with those from the older well-established Thurstone and Chave and Likert tests shows a considerable degree of correspondence. The inference that can be made is that whatever is measured by the older tests is measured approximately as well by the self-rating scale. Furthermore, the self-rating scale was easy to construct, straightforward to administer and simple to score.

Given then a measuring instrument of this nature, some examination of the theoretical basis on which it rests is necessary. Initially, certain assumptions were made about attitude scaling in general. Briefly, they involved the following points: First, that if a simple technique could be substituted for a complicated one, more practical advantage would be likely to accrue. Second, that an individual *can* make an effective assessment of his own attitude—at least towards some 'psychological objects.' Third, that an attitude-expressed-in-action reflects essentially the result of a choice made by the individual from a number of related alternative modes of action. Fourthly, for any individual, *intensity* of attitude is dependent on a number of highly personalized factors and that consequently, in some situations, he is his own best judge of the relative strength of his own attitude. Fifthly, if this last point is correct, a test requiring him to discriminate among his own attitudes is more likely to return results that will coincide with his behaviour (his attitude-in-action) than one that seeks to discern the intensity of his attitude by reference to the generalised responses of other individuals. In other words no assumption was made here about the distribution of attitudes (normally or otherwise) throughout the parameter.

Within the test itself are to be found further assumptions, namely, that a respondent *can* conceptualise his own attitude in numerical terms, and that a numerical interpretation conveniently avoids some of the pitfalls to which verbal interpretations are prone; in particular, confusion in terminology. Furthermore, it was assumed that the very nature of the numerical response required would constrain the respondent to make an internal comparison, as it were, of relative attitude strengths, and that because reference to specific circumstances or situations was avoided he would be able to do this in general terms for each attitude area tested.

The author was originally prepared to hypothesise that some such phenomena as a general attitudinal predisposition existed in each invididual. Some people, obviously enough, tend to be generally enthusiastic about a great many things, others to be correspondingly unenthusiastic, while others again have major and minor dominant and subordinate likes and dislikes. Each individual has then a relatively unique configuration of attitudes that would require any examination to be from the particular rather than the collective standpoint. While such a statement could be taken to lend support to Doob's (1954) argument that attitudes should be subsumed under general personality theory this is another issue, although it is not without relevance to the approach adopted here. The basic concern, however, is with the evaluation of a measuring instrument designed for use in a practical situation.

VI.—CONCLUSION.

The self-rating scale under consideration did make it possible to relate attitudes recorded by one individual to other attitudes recorded by the same individual. Furthermore, there appeared to be, understandably enough, a degree of correspondence between individual and group results. Courses of study ranked highly by individuals tended to receive high scores also when individual scores were aggregated.

It seems reasonable in the prevailing situation to use the tests to measure changes in attitudes during the teacher training period, to discern attitude idiosyncracies reflected in individual scores so that appropriate action can be taken, to provide a further aid for student guidance within the training programme and to help arrive at a basis for predicting attainment, application and future teaching competence. There is every reason to believe that the technique will have wider application than the restricted field of the initial experiment. It is, for example, being used in connection with an investigation into children's attitudes towards arithmetic in primary schools.

VII.—References.

Doob, L. (1954). The behaviour of attitudes. *Psychol. Rev.*

Klineberg, Otto (1954). *Social Psychology*. U.S.A.: Henry Holt.

Likert, R. (1932). A technique for the measurement of attitudes. *Arch. Psychol.*

Thurstone, L., and Chave, E. (1929). *The Measurement of Attitudes*. Chicago: Univ. Chicago Press.

(*Manuscript received* 21*st December*, 1961.)

SCHOOL PHOBIA

By MAURICE CHAZAN

(Department of Education, University College of Swansea)

SUMMARY. This paper analyses thirty-three cases of school phobia referred to a child guidance centre between 1949 and 1959. The findings agree with those of previous writers with regard to the level of intelligence, home background, and personality features of such children, but suggest also that educational factors have been underestimated in the consideration of both aetiology and treatment.

Most of the children were of at least average intelligence, came from materially good homes, and were members of small united families, where the mother-child relationship was very close. About two-thirds of the group showed marked dependency and emotional immaturity. Traumatic events at home featured rarely in the case-histories, but in twelve homes at least one parent was chronically ill, physically or mentally. The majority of the children were in good health, but in six cases refusal to go to school followed immediately after a period of illness or hospitalisation.

Investigation of their school life and attainments suggested that factors such as educational retardation and the pressures of a new and more exacting school situation were often of relevance. Feelings of social inferiority seemed to contribute to the breakdown of four out of the eight children referred from grammar schools.

Methods of treatment included individual psychotherapy, group methods of remedial teaching or play therapy, and placing the child in a less demanding school situation.

I.—INTRODUCTION.

CHILDREN who refuse to go to school do not form a homogeneous group. School phobia has, however, been described as a reasonably well-defined syndrome (Broadwin, 1932; Johnson, *et al.*, 1941), which can, to some extent, be distinguished from truancy (Warren, 1948; Hersov, 1960), although there may be overlapping features. The child persistently refuses to attend school and shows the symptoms of an anxiety state, but does not attempt to absent himself from home, or to conceal the situation from his parents. A panic state arises if the child makes an effort, or is urged, to go to school. The parents usually do everything possible to persuade the child to attend school, are deeply troubled by their lack of success, and are ashamed because the child has come to the notice of the school attendance authorities. In the case of the truant, on the other hand, parental supervision is often lax and sometimes the child's absence from school is even condoned by the parents (Tyerman, 1958).

While there is evidence that an increasing number of cases of school phobia is being referred to clinics,* the incidence appears to be very small. Few children go so far as to refuse to attend school, however troubled they may be with anxieties in connection with work or social life there. Although Nursten's (1958) cases constituted 8 per cent. of the clinical population during a two-year period, other studies (Model and Shepheard, 1958; Morgan, 1959) report that school phobia cases form about 2–3 per cent. of referrals. At the Liverpool Education Committee Child Guidance Centre, only thirty-three cases of school phobia (twenty girls; thirteen boys) were seen during the period April, 1949, to December, 1959, inclusive; this represents just over 1 per cent. of the total number of referrals to the Centre. Eight of the cases were referred between 1949 and 1953, fifteen between 1954 and 1958, and ten during 1959.

---

* *Lancet*, 30th January, 1960.

This paper describes the children and their family backgrounds, the methods used to deal with the problem, and the outcome of treatment. Each case was dealt with by the full clinic team of psychiatrist, psychologist, and social worker; in addition to group therapy remedial teaching was available as a form of treatment.

II.—PREVIOUS WORK.

Previous studies, both American and British, stress the fundamental importance of the mother's role in the aetiology of school phobia, and tend to discount scholastic factors (Johnson, 1941; Model, 1958; Kahn and Nursten, 1958; Morgan, 1959). The mother is described as over-protective and closely identifying herself with the child, who becomes over-dependent on her. Morgan (1959) and Hersov (1960, b) found a change of school or class to be an important precipitating factor, but intellectual or educational retardation are not considered relevant factors. Green (N.A.M.H. Conference Report, 1959) has indicated that an adjustment in the school situation may be a considerable help in solving the problem, but is in agreement with reports (Warren, 1948, and Hersov, 1960 a) that school phobics, unlike many truants, are usually quite able to cope with their school work.

In general, the emphasis in treatment has been placed on therapy for the mother and child (Kahn, 1958; Davidson, 1961), other methods being considered as likely to fail. The need for close co-ordination between clinics and other agencies has been stressed (Morgan, 1959).

III.—DESCRIPTION OF THE CHILDREN.

(i) *Age at time of referral.* The age range of the thirty-three cases was 5 years 5 months to 14 years 2 months, the average age being 10 years 6 months. Fifteen children were under 11, and eighteen over 11 years of age. As in Morgan's group (1959), there was a peak in the age range 11 to 13 years (twelve cases). Nine cases were between 11 and 12 years old. Surprisingly, there was only one child in the 5 to 6 year old group, and only one child who was immediately concerned with the selection examination at 11.

(ii) *Type of school attended.* Sixteen children attended primary schools, of whom twelve were in the junior department. Seventeen attended secondary schools, eight of these being at grammar schools. One boy was at a technical school.

(iii) *Stage of school career.* Nineteen of the children were in the first year of their school (infants, junior, or secondary), three were in their second, nine in their third, and two in their fourth year. Eleven were in their first term of the first year of a new school, and as many as twenty were in their first term of a new class or school.

These figures support Morgan's (1959) and Hersov's (1960, b) view that a change of school or class must be considered an important factor in the break with school, though Burns (N.A.M.H. Report, 1959) did not find an actual change of school to be such a frequent factor.

(iv) *Early difficulties.* In seven cases there had been difficulties in connection with birth or pregnancy, and in two cases difficulties in connection with feeding.

(v) *Physical condition.* If one excludes psychosomatic symptoms specifically produced by the thought of school or attempt to go to school, twenty-one cases were in good general health at the time of referral, and had not suffered from ill-health for some time previously. Of the others, four suffered from asthma ; two had epilepsy ; and six were in poor health, suffering from frequent colds, bronchitis or tonsillitis.

Ten cases had a history of hospitalisation, but only three had been in hospital before the age of five.

Two children were described as physically immature ; one had a congenital physical handicap ; and one was sensitive about his obesity.

It has been suggested (Estes, *et al.*, 1956) that school phobia is sometimes the result of illness, e.g., influenza, which is often followed by depression. In six cases, refusal to go to school did follow immediately after a period of illness or hospitalisation.

Twenty-one children complained of ailments which were clearly psychogenic and related to the school phobia.

(vi) *Level of intelligence.* All the children were tested on the Terman-Merrill Scale (Form L) ; the Roberts-Mellone adjustment was applied to the I.Qs. obtained. The I.Q. range was 74–132, the mean I.Q. of the whole group being 105·8 (with a standard deviation of 14·9). There were ten cases with I.Qs below 100, eight with I.Qs. between 100 and 109, and fifteen with I.Qs. above 110. The mean I.Q. of the eight grammar school pupils was 121. With two exceptions, the group as a whole does not consist of *very* bright children, but it is above the general level of the clinic's population (average I.Q. approximately 95, if cases specifically referred for backwardness are excluded), and there were only four dull children in the group.

Previous studies have found a fairly similar range of intelligence quotients.

(vii) *Extent of educational difficulties.* The basic educational level of the children attending primary and secondary modern schools was assessed by means of the Schonell Word Recognition Test, Vernon's Graded Arithmetic-Mathematics Test, and a school report. There were twenty-four children in this group, of whom one had just started school ; six had no educational difficulties ; four were slightly backward, but able to read adequately and to cope with the work of their class ; but as many as thirteen (mean I.Q. 98·7) were very backward, i.e., they were experiencing great difficulty with their work, and were in need of special educational treatment.

It is of interest to note that the educational attainments of this group, consisting mainly of highly anxious children, do not seem to lend support to the hypothesis (Lynn, 1957) that anxious children tend to be better at reading than arithmetic. Reading age was above arithmetic age by a year or more in ten cases, but in seven cases, arithmetic age exceeded reading age by at least a year, and in fifteen cases there was no significant disparity between attainment levels in reading and arithmetic. Of the thirteen very backward children, no more than four were very backward in number.

Of the nine selective secondary school children, only one had difficulty with the work of his class.

These facts suggest that while failure to cope with school work is not an important factor in the grammar or technical school cases, educational retardation cannot be ruled out as playing a part in the causation of school phobia in about a half of the other children, i.e., those at primary or secondary modern schools. It cannot be assumed that the standard of work of these children is ' invariably good,' as stated, for example, by Hersov (1960, b).

(viii) *Personality features.* Previous studies have stressed that school phobics (i) show over-dependency, (ii) exhibit symptoms of generalized anxiety, (iii) are not aggressive or delinquent, but immature, over-protected, prone to temper tantrums, sleeping difficulties, and timidity. These features commonly occurred in this group. Apart from two children, who were emotionally blunted, the group showed a good deal of affect and anxiety, though not a few were overtly cheerful as long as they did not have to go to school. Two-thirds of the group showed marked emotional immaturity and dependency, timidity and nervousness. Twelve presented behaviour difficulties at home, but not one was reported to be difficult at school, and only one had any delinquent tendencies. Davidson (1961) stresses the frequent occurrence of depression in her series of cases, but, in the Liverpool group, as in Hersov's (1960, b) anxiety symptoms featured more frequently than depressive reactions.

IV.—HOME CONDITIONS AND FAMILY CIRCUMSTANCES.

(i) *Position in the family.* Previous studies agree that the children tend to be only children or extreme in birth rank, and that this is probably of significance in relation to the mother's handling of the child. In this group, there were as many as twenty-one who were extreme in birth rank (fourteen youngest, seven eldest) and there were six only children. The other six children were all the middle of three. There was one adopted child, but there were no illegitimate children, nor any foster-children in the group.

(ii) *Size of family.* Only three of the group came from families with more than three children. Marked sibling rivalry was noted in only four cases, and in no more than two instances did the trouble follow the birth of a sibling.

(iii) *Socio-economic status of the home.* The fathers were mostly workers earning reasonably good wages, in skilled or semi-skilled work (twenty-two out of the twenty-eight fathers). Two had their own business, one was a seaman, one had retired, and two were unemployed.

The material standard of the home was unsatisfactory only in four cases, chiefly owing to the structural condition of the house. Most of the homes were described as spotless and beautifully kept.

Financial strain was apparent in only one case, where both parents were in poor health. Twenty-one of the mothers were not working at all; five were working full-time, including one at home; and seven did part-time work—so that only four of the mothers were away from home all day.

Previous studies report that all social classes are represented, and that the children come from materially good home backgrounds.

(iv) *Broken homes.* Burns (N.A.M.H. Report, 1959) found few badly broken homes or severe traumata in the background of children seen in Birmingham for school phobia.

In the Liverpool group, the mothers were all active members of the family, and in only three cases had the child been separated from the parents before the age of five. Five homes lacked a father (three because of divorce or separation, two because of death), and one father was away at sea for long periods

Moves of house had been rare. The families had mostly lived in their homes for a long time and in some cases were a close 'in-group.' In only one case had frequent changes of accommodation contributed to the insecurity of the child.

(v) *The mother-child relationship.* On the whole, the mothers were intelligent, co-operative and capable of good insight, but weak, anxious, over-protective and worriers by nature. There was only one case where there was any parental neglect ; this boy is the sole delinquent of the group. In twenty-one cases the mother-child relationship was over-close, grossly so in twelve cases ; but in only three cases did the problem seem to be entirely due to the child's fear of leaving his mother. In five cases, the close mother-child relationship was reinforced by indulgent grandparents and in several others by the absence of a father or by the father's antagonism to the mother or child.

Morgan (1959) comments that the mothers of school phobics tend to be elderly and Burns (N.A.M.H. Report, 1959) considers that they are nearly all neurotic in some sense.

(vi) *The father-child relationship*. It is not so easy to assess the attitudes of the fathers, because they are not often seen by the visiting social worker or at the clinic. They appear to play a rather passive role in the family, though the majority were described as good fathers, patient and sensible. Six were weak, indulgent and anxious ; four were said to be domineering and bad-tempered, but only one took little interest in the child. Previous studies describe the fathers as ineffectual, unauthoritative, or uninterested.

(vii) *Husband-wife relationship*. Burns (N.A.M.H. Report, 1959) reports that, in as many as a half of his families, there were strained relationships between husband and wife, but, in the present group, in only four of the twenty-eight homes where both parents were present was the parental relationship admitted to be unsatisfactory. It is, of course, far from easy to elicit reliable information about parental disharmony.

(viii) *State of health of parents*. There was chronic ill-health, physical or mental, in twelve homes, in two of these both parents being ill.

V.—ALLEGED PRECIPITATING FACTORS.

The children are usually very reticent about the reasons for their refusals to go to school, even if they are aware of them. They are prone to rationalization and projection, and so too much importance cannot be attached to reasons given by the child, though these should not be ignored, as they may provide clues to vital factors and point to adjustments which need to be made in the school or home situation.

In twenty-two cases difficulties in the school situation were alleged ; in four cases there were traumatic incidents at home ; in six cases a period of illness seemed to be responsible for the final break with school, but in four cases no upset of any kind came to light.

The difficulties in the school situation allegedly arose from a dislike of being punished or shouted at in class (eleven cases) or being afraid of other children (six cases). Other complaints (in two cases each) were : dislike of a particular lesson, sex play, having to stay for school meals, or a traumatic event in school (e.g., being locked in a classroom accidentally) ; and (in one case each) fear of being in a confined space, fear of crowds, anxiety over an examination, or fear of an epileptic attack occurring in class.

In the grammar school cases, the refusal to go to school arose mainly as a result of their failure to adjust to strange surroundings and new demands. Four of the children were from homes with no grammar school tradition, and felt a sense of social inferiority (Chazan, 1959). Two of the boys (one of whom was depressed about his adverse school reports) were showing signs of incipient psychosis.

The upsetting incidents at home included the birth of a sibling, the mother having an accident, and the loss of a near relative.

VI.—TREATMENT.

The following treatment measures were recommended, in addition to advice to parents and contact with the schools:

Change of school	19 cases.
Individual psychotherapy	16 cases.
Remedial teaching at Centre	9 cases.
Group therapy	9 cases.
Sedatives	5 cases.
Hospital treatment	2 cases.

A change of school, which did not preclude psychotherapy or remedial teaching at the Centre, was recommended on the grounds of (i) the child's fixed hostility to his present school, (ii) a considerable degree of educational or intellectual retardation, or (iii) the child being too immature, or in too anxious a state, to cope with the life of his school. The changes recommended were to these types of schools:

Day Open-Air	7 cases.
Residential School for Maladjusted Boys	4 cases.
Secondary Modern School (from grammar school)	4 cases.
Residential Open-Air	3 cases.
Day School for E.S.N. pupils	1 case.

It will be noted that the changes recommended were to schools which would offer less pressure than the type of school which the child was resisting, in the sense that there would be smaller classes, shorter hours, less robust companions, or less emphasis on academic attainment. A mere change of school on the same level was not considered likely to be successful in any of the cases seen.

A number of the children attended the Centre for remedial teaching or group play therapy (or both), sometimes on a daily basis. The availability of these forms of treatment meant that (*a*) immediate action was taken, (*b*) the child was helped to separate from the mother, (*c*) contact was maintained with school work and with other children, (*d*) attention was paid to educational weaknesses where necessary, (*e*) the child was not just kept at home while awaiting a vacancy at a special school.

The parents were advised not to use force, and there was no recourse to legal procedures. No child was taken into the care of the Local Authority.

The return to the child's own school on a full-time basis was in some cases preceded by his going part-time for a while, with the co-operation of the head-teacher. In one case, the child seemed to begin to overcome her fears by going to see her teacher informally at the end of the school day. In two most difficult cases, the offering of the alternative of going to a residential or a day school was used as a last resort.

Two of the children went back to school spontaneously; in two cases there was a lack of co-operation with the Centre.

VII.—OUTCOME OF TREATMENT.

Most of the cases were followed up for some considerable time. Twenty-nine of the children were back in school within 5 months of their referral (nineteen within 2 months). Only one boy—the technical school pupil, who had a grossly abnormal cerebral dysrhythmia—refused to return to school life, and later also refused to go to work.

Twenty-two of the children adjusted satisfactorily in primary or secondary modern schools, including (*a*) seven children who returned after a period at a special school and (*b*) four of the grammar school pupils, who adjusted immediately on their transfer to a secondary modern school.

Seven children settled down well in special schools, presenting no difficulty there.

The resolution of the children's conflicts with regard to school attendance seemed to help them to make better progress with their school work, and to mature in personality.

VIII.—DISCUSSION.

It is interesting to note to what an extent there are common features in the various groups studied. Poor material circumstances and low intelligence play little part in the causation of school phobia, while over-protective parental attitudes, the effect on the family of serious mental or physical illness in one or both parents, and the demands of a new situation in school seem to be important predisposing or precipitating factors. It is possible, too, that in some cases disturbance results from a period of illness in the child. Many of the children experience little or no difficulty in coping with school work. There is evidence, however, to suggest that educational retardation should not be discounted as a factor in the causation of some cases of school phobia. It is likely that the more intelligent a child is, the more aware he is of his backwardness and the more sensitive he becomes about it. It is agreed that psychotherapy is often necessary, but whatever other treatment measures are implemented, it seems desirable that remedial teaching should be given to the child who is educationally retarded, unless he is too disturbed to co-operate.

Group play therapy, too, might be considered as offering a contribution to the solution of the problem, in that the child may be helped to maintain contact with other children while making a fresh relationship with an adult other than his mother.

A change of school should not be regarded as necessarily a ' stop-gap', ' face-saver,' or likely to fail (Kahn, 1958). It is not surprising that there should be a small number of children who require a form of school life which makes only limited demands on them. Just as the ravages of illness may leave a child too delicate to meet the physical demands of the ordinary school, so nature and nurture sometimes produce a timid, anxious and immature child who is unable to cope with the normal stresses and strains of school life. The children who were transferred to other types of school quickly settled down, making no fuss about going to schools some distance away from their homes, even travelling by bus and staying all day. The increased amount of individual attention, made possible by the smaller classes of the special school, undoubtedly helps the child to readjust to school life, and the wider provision of day special schools for maladjusted pupils should help a number of difficult cases of school phobia.

It is important that, when a change of school takes place, the problems of the child are fully understood by the staff of the new school, and the psychologist has a valuable role in establishing and maintaining liaison between clinic and school.

Social factors appear to be relevant in a few cases, particularly with regard to some of the grammar school children, whose reluctance to go to school was an unwillingness to leave the neighbourhood and to lose sight of old associates rather than a fear of leaving home. These children, who felt uncomfortable in the grammar school, presented no difficulties at all as soon as they were in a secondary modern school. While it is a pity that these children will not take the educational opportunities offered to them, and while every effort should be made to encourage the bright pupils to continue in a grammar school, little can be done in face of a child's hostility to grammar school life.

ACKNOWLEDGEMENTS.—The writer would like to thank Dr. G. Stuart Robertson, Deputy Principal School Medical Officer, Liverpool, as well as Dr. H. S. Bryan, Dr. I. Leveson, and Dr. P. Pinkerton, Consultant Psychiatrists, for their kind permission to publish the data of this study.

IX.—REFERENCES.

BROADWIN, I. T. (1932). A contribution to the study of truancy. *American J. Orthopsychiat.*, 2, 253-259.

CHAZAN, M. (1959). Maladjusted children in grammar schools. *Brit. J. Educ. Psychol.*, 29, 198-206.

DAVIDSON, S. (1961). School phobia as a manifestation of family disturbance: its structure and treatment. *J. of Child Psychology and Psychiatry*, 1, 4, 270-287.

ESTES, H. R., HAYLETT, C. H., and JOHNSON, A. M. (1956). Separation anxiety. *Amer. J. Psychotherapy*, 10, 682-695.

HERSOV, L. A. (1960, a). Persistent non-attendance at school. *J. of Child Psychology and Psychiatry*, 1, 2, 130-136.

HERSOV, L. A. (1960, b). Refusal to go to school. *J. of Child Psychology and Psychiatry*, 1, 2, 137-145.

JOHNSON, A. M., FALSTEIN, E. I., SZUREK, S., and SVENDSEN, M. (1941). School phobia: a discussion. *Amer. J. Orthopsychiat.*, 11, 702-711.

KAHN, J. H. (1958). School refusal. *The Medical Officer*, 100, 337-340.

LYNN, R. (1957). Temperamental characteristics related to disparity of attainment in reading and arithmetic. *Brit. J. Educ. Psychol.*, 27, 62-67.

MODEL, A., and SHEPHEARD, E. (1958). The child who refuses to go to school. *The Medical Officer*. 100, 39-41.

MORGAN, G. A. V. (1959). Children who refuse to go to school. *The Medical Officer*, 102, 221-224.

NATIONAL ASSOCIATION FOR MENTAL HEALTH (1959). Truancy—or school phobia? Proceedings of the 15th Inter-Clinic Conference.

NURSTEN, J. P. (1958). The background of children with school phobia. *The Medical Officer*, 100, 340-342.

TYERMAN, M. J. (1958). A research into truancy. *Brit. J. Educ. Psychol.*, 28, 217-225.

WARREN, W. (1948). Acute neurotic breakdown in children with refusal to go to school. *Archiv. Dis. Child.*, 23, 266-272.

X.—APPENDIX.

ILLUSTRATIVE CASE STUDIES.

A.B. (a girl of 11, I.Q. 121) was referred to the Centre shortly after beginning her first term at a grammar school. She refused to attend school, becoming hysterical at the idea of going back. She had never presented any problems previously; she was in good health, and had never been in hospital. Emotionally, she was immature, dependent, anxious and tense. She said that she missed her old school friends, and alleged that a teacher at her new school had shouted at her. She was an only child; her father had died suddenly five years previously. He mother was anxious, but had good insight. She worked part-time, always being at home when the girl returned from school. The home was rather drab and in a poor district, but clean and comfortable. The girl refused a transfer to another grammar school, insisting on going to her local secondary modern school. This was arranged, and she immediately settled down happily.

C D. (a boyd of 7, I Q. 120) was referred just after commencing his first term at a junior school. There had been some previous reluctance to go to school while he was in the infants' department, and he had had a long series of illnesses during his first year at school; now he was adamant in his refusal to attend. He said that he did not like his teacher, and that the other boys hit him " for the slightest thing." He had difficulty in mixing with other children and was timid, apprehensive, and insecure. In spite of his intelligence, he could not read at all, and was unable to cope with the work of a junior school.

There had been many changes of residence; the family had just moved into their first house after thirteen changes of apartments. They set themselves a very high standard, and the house was beautifully kept. The boy was the middle of three children. The father was somewhat short-tempered, and the mother was rather weak and over-anxious.

For a period the boy attended the Centre almost daily for remedial teaching, play therapy, and psychiatric interviews. His mother was also seen frequently. He was transferred, three months later, to a day open-air school, where his attendance was excellent. After five terms there, he was able to return to a junior school (this was a different one, to give him a fresh start) and later was transferred without difficulty to a secondary school. Supervision at the Centre was maintained throughout this period, and the necessary liaison with the schools was established by the psychologist.

E.F. (a boy of 13, I.Q. 115) refused to attend school during his second year at a secondary technical school. He was the elder of two boys, with an over-anxious, inconsistent mother. He complained of symptoms of depersonalization; he felt ' far away ' at times while in class, and began to be afraid to be away from home. The home was very clean and tidy, and there were no adverse material circumstances. The boy was not experiencing any serious educational difficulties.

Investigations at hospital, including several E.E.G. readings, suggested an irritable focus in the upper part of the right occipital lobe. The boy was treated in hospital, but could not be persuaded to return to school, nor did he go to work when he reached the appropriate age. His mother was not consistent in her co-operation with suggested measures for his rehabilitation.

(*Manuscript received* 15*th November*, 1961.)

SYMPOSIUM: EDUCATIONAL RESEARCH TODAY

III.—The Tools of Research—Men and Machines

By STEPHEN WISEMAN
(University of Manchester)

" In view of the very large sums of money that are spent on education every year, the expenditure on educational research can only be regarded as pitiable." Thus said the Crowther Report in 1959 (para. 697). There is no doubt that it is true: roughly 0·02 per cent. of the total national expenditure on education goes towards research. Many people, faced with this fact, assume that the major task before us is to secure a massive increase in financial support for research: if this is done, all will be well. But is this really so? What would happen if some fairy godmother suddenly produced a million pounds for educational research? Could this sum—or even one-tenth of it—be used profitably and productively? It is one of the contentions of this paper that it could not, and that the provision of really strong financial support for research would cause most university departments and institutes of education extreme embarrassment. This is because of the shortage of men and machines—and the organisation necessary to employ these effectively.

Research workers must be trained before they can be employed as such, a fact that is not recognised by many teachers, administrators and civil servants. The usual method of training, as in the other sciences, is through the medium of the master's degree. Since the war there has been a considerable increase in the number of candidates for M.Ed. and M.A.(Ed.), at British universities. It should be noted, however, that this increase is very unevenly spread: in many universities the numbers graduating are extremely small. And it must also be remembered that many candidates present theses in fields other than educational psychology.* Even those who present theses on topics in educational psychology cannot all, with certainty, be judged as trained in research methods. The amount, and kind, of training given to such students varies very much from university to university, and for some students in some universities it consists of little more than learning how to calculate a standard deviation.

I regard the M.Ed. and the M.A.(Ed.) as *training* degrees. Their value lies in their effect on the student rather than in the end-product of the research. It is true that many M.Ed. researches produce valuable results, and that, over the years, the total contribution of M.Ed. students in an active department may be considerable. But the kind of topics which an individual student can tackle for his first investigation are extremely circumscribed, both in scope and in sophistication of method. If follows that we would expect M.Ed. theses to contribute only marginally to the real research output of a university department. It is sometimes urged that by careful planning and research direction, the efforts of individual students can be co-ordinated so that, over a period, major advances can be made in certain well-defined areas of the field. This is a persuasive argument, but a dangerous one. The selection of a research topic by the student himself is an important part of his training (and also, in my view, an important element in the selection of students for admission). Presenting a student with a ready-made topic (and, all to often, a ready-made method) may be all very well for the young post-graduate physicist or chemist, but it is a different matter for the experienced teachers who form the bulk of our M.Ed. applicants. They have—or ought to have—their own problems that they wish to investigate, problems already familar to them as teachers, and ones which provide strong motivation. To switch them to something entirely different is a decision not to be taken lightly. The best way to achieve some measure of co-ordination is the ' natural ' way: the attraction of students to particular teachers by virtue of the reputation these teachers have gained in particular sections of the research field.

---

* I am not suggesting that research in, say, the history of education is unimportant, or that theses in educational philosophy should be proscribed. But it seems to me that productive research in education, focussed on current problems, must inevitably be heavily biassed towards educational psychology, using the methods and techniques developed in this field. Even in the growing field of sociology of education, I am convinced that a basic training in sampling theory, multiple regression, factor analysis and the analysis of variance is a desirable pre-requisite for research workers.

In parenthesis—and with reference to the co-ordination of research—some comment is appropriate on the research groups of teachers and training college lecturers described in Mr. Burroughs' paper. Here I would like to differentiate between *research* and *enquiry*, to distinguish between ' hard ' research, carried out by trained personnel and using all the statistical and methodological techniques available and appropriate, and those enquiries that are capable (at any rate in the initial stage) of attack by teachers and lecturers experienced in education but not in research methods. I am concerned only with ' hard ' research in this paper. I recognise that there is no clear demarcation between the two categories I have delineated, but the contrast is a real one. I do not wish to suggest that the formation of groups of teachers and training college lecturers for the purpose of group enquiry is unimportant and unproductive. It has two very clear advantages: first, to the people taking part, as a valuable education in itself, and as a stimulus to further study, leading some to more systematic training; second, the end-products of such enquiries may form the basis for more penetrating research. But I suggest that the research productivity of any department or institute should be judged on its output of *research*, and that *enquiry* be judged by its success in supporting this, in stimulating serious students and in producing ' ground-clearing ' studies.

By means of the master's degree we succeed in training a relatively small number of research workers each year. Once we have trained them, they are now fitted to engage in serious research. Some, of course, are more ' fitted ' than others, and many do not wish to go any further. Nevertheless, there remains a number of very able students who are eager to go further. What happens to them? One or two might be lucky enough to be appointed to a lectureship in a university department where they will have opportunities and facilities (varying quite markedly in amount from one university to another) for further research. But the majority will remain in their schools, and look hopefully to the university to permit them to continue research—part-time, as with M.Ed.—for the Ph.D. degree. Most of these will find themselves disappointed, since Ph.D. regulations, framed with the needs of the pure scientist in mind, are likely, with their attendance requirements, to prevent them from pursuing this course. I have already, some years ago, pointed out* this nonsensical barrier to the use of our trained research workers, but my plea seems to have found little support. I suspect that this is largely caused by the general climate of opinion in universities and the attitude towards education as an academic study. Indeed, I regard this as one of the root causes of the ' pitiable ' amount spent on educational research. Many of those concerned with education, and particularly higher education, think of departments of education as places for the peddling of hints and tips on classroom techniques, for the arrangement of teaching practice, and for the provision of leisure time for students recovering from the hard intellectual grind of their degree courses. The idea of education as an academic study and an intellectual discipline (a word much in favour—and with a wider connotation—with those who have had no contact with educational psychology), let alone as a field for serious quantitative research, is quite foreign to them.

---

* Higher degrees in Education in British universities. *Brit. J. Educ. Studies*, 1953, 2, 54-66.

Consider the evidence. There are, in England and Wales, thirty-nine professors of education. Just as professors of physics are distributed, in their specialities, over the sub-fields of mathematical physics, atomic physics, theoretical physics, and so-on, so we might expect out professors of education to be educational psychologists, historians, sociologists, philosophers and comparative educationists. And we might judge that educational psychology ought to be at least as strong in numbers as any other section. (Many of us might see, as a reasonable distribution of ten professors of education, five psychologists, two historians, and one of each of the others. But, perhaps this is tendentious* !) What do we find, in fact ? Of our thirty-nine professors, eight are members of the British Psychological Society, and of these eight, five are Fellows (Birmingham, London, Manchester (two) and Nottingham). It must not be assumed that this very small number of psychologists is caused by the overwhelming popularity of historians, philosophers and comparative educationists : the largest single category is one that can best be described as educationists without prefix. Many in this category have publications of a critical and scholarly kind to their credit, but not all can legitimately be said to have engaged in educational research of any kind.

A university teacher in any other subject is expected—indeed, required—not only to teach, but to carry out research. When such a teacher is appointed, the university looks for evidence of scholarship and of relevant experience, and also looks—very searchingly—for evidence of research and publication. Lecturers in education, since they are to work in a professional department, must usually show, in addition, evidence of success or even distinction in teaching in schools. A professor of education, then, might be expected to have distinguished himself in the outside world of education, and also to have made some weighty contributions to research and theory. The facts show that this is not universally so. Clearly a chair of education in some places is of a different shape, and has a different function, from chairs in other departments and faculties.

This state of affairs ought to cause us grave concern, for two reasons. First, and most obviously, for the evidence it gives of the deplorable ' image ' of education as a subject in many universities. (Remember that most professors are appointed by committees of university senates.) Secondly, because of the inevitable effect it must have on the kind, and amount, of research carried on in education departments, and on the number of research workers trained there. It might be thought that the background and experience of the professor is not necessarily so important, as long as his staff contains lecturers expert in the research field. But lecturers who are interested in doing research in educational psychology, and in teaching its methods, are unlikely to be attracted to such departments. And the professor's own interests, and cast of mind, influence very profoundly the kind of appointments that are made, as well as the facilities provided for research. In addition, progressive developments within a department, such as the institution of advanced courses and new diplomas, tend to reflect the major interests of the head of the department. Such advanced courses may provide important stimulation to research, and act as breeding grounds for research students.*

---

* It should be noted, however, that my 1953 analysis of the fields of degree theses in British universities showed 55 per cent. of psychological topics in the decade 1938-47, and 68 per cent. from 1948 to 1951. (*loc. cit.*, Table 3).

It may be pointed out that many institutes of education, and some departments, appoint research fellows, and that this enables much important research to be done. In many cases, this appears to be the way—almost the only way—in which a responsibility to carry out research is discharged, and the discharge made evident. It is easier to appoint a succession of temporary research fellows than to build up a teaching force which is adqeuately research-minded and research-qualified. In some cases the fellow is appointed with an almost free hand in the choice of his research field: the possibility of building up a coherent research pattern and plan is deliberately discarded. Although it is recognised that one or two of these research fellows have produced interesting and important work, it is also true that the total output from this source over the last decade is almost insignificant in its impact on theory and practice. These one-year miscellaneous fellowships are often little more than a sabbatical holiday for their holders.

There is one way in which this sombre and depressing picture could be radically changed, and a firm base provided from which educational research could grow to reasonable proportions. That way is the institution of more chairs in education, including *research chairs*. We know from experience in other subjects (and from American experience in education) how a research professor, largely free from the administrative load which so hampers his colleagues, can build up a strong research department which can not only tackle long-term, co-ordinated research projects, but also give a thorough training in methods and techniques to a growing number of students. To those who fear, with some justification, too radical a divorce of teaching from research, the alternative is the creation of new teaching chairs, particularly in Educational Psychology. This would permit the spreading of the load of administration, and give more time for the active direction of research by senior members of staff. Only with such radical innovations can we hope to make any significant advance from the present position. Research development depends primarily upon personnel. At present the number of people in this country adequately trained to pursue (let alone direct) ' hard ' research is pathetically small—and I doubt whether it is growing at any significant rate. Many of those that we do train, we lose. And lecturers in education see no career prospects in specialising in educational psychology and research: the prospect of appointment to a chair is so remote as to be almost chimerical. But with the establishment of research chairs and chairs in Educational Psychology, there would be a clear path through an increased number of fellowships, senior lectureships and readerships. And the stimulation to the day-to-day activities of a department and institute of education produced by the existence of a research chair must not be overlooked nor under-estimated.

---

* I am thinking particularly of such courses as those at Birmingham (Post-graduate Diploma in Educational Psychology, Diploma in the Psychology of Childhood, Diploma for Teachers of Educationally Sub-normal Children) and at Manchester (Diploma in Educational Guidance, Diploma in Educational Psychology (Child Guidance), Diploma in the Education of Handicapped Children).

Now let us turn from the problem of the supply of human material to the provision of mechanical aids to research. Like other conditions in departments and institutes, the position here seems to vary enormously over the country. I have heard of one department of education where lecturers on the staff have to borrow an electric desk calculating machine from another department if they have any heavy computations to do. The provision of at least one such machine would seem to me an irreducible minimum in the way of mechanical aids—and a machine specifically chosen for its ease in accumulating squares (e.g., Friden, Marchant or Madas).

The advent of the electronic computer has revolutionised the outlook for research, and brought within our grasp research projects which, a few years ago, would have been far beyond our resources in time, labour and personnel. Its impact on educational psychology has been (or ought to be) such greater than on many other sciences, because of our interest in, and dependence on, correlational analysis leading to multiple regression and factor analysis. With the electronic computer we can now consider with equanimity dealing with correlation matrices covering thirty or forty or more variables. But we cannot hope to exploit this powerful aid unless we have personnel trained to use it. The number of universities without an electronic computer—or access to one—is rapidly growing smaller : but how many of the education departments of those possessing a computer are able to take advantage of its presence ? It is of little use to regard the computer as an outside ' consultant,' to be employed by laymen. Programmes for correlations and for analysis of variance might, if one is lucky, already be set up at the instigation of other departments ; programmes for multiple regression, for factor analysis and for factor rotation are extremely unlikely to be available. To have to rely upon the computer staff to punch data on tape is likely to delay one's research by many weeks, if not to produce a blank refusal. It is essential to have, within the department itself, clerical and computing staff able to punch tapes (and to possess the necessary machinery), and most desirable to have at least one member of staff skilled enough to write the required programmes. To those who regard this as impossibly utopian, may I assure them that we have achieved all this in Manchester over the last few years, with very little difficulty.

The use of a computer brings new problems, of course, and particularly with regard to its use by research students. The computer widens the horizon of the M.Ed. student, and brings within his grasp problems of a range and complexity which previously were quite impossible. But it is necessary to be sure that the student *knows* the techniques he is using, and appreciates their strengths and limitations. No-one should be permitted to use the computer for the calculation of correlations, for example, unless he has demonstrated his ability to calculate these ' by hand ' and to interpret them.

May I attempt to summarise the main burden of my argument ? The significant expansion of educational research is, I suggest, impossible at present, mainly because of the shortage of trained research workers. When we complain about the microscopic amount spent on research, let us reflect on the possibility that we ourselves are mainly responsible for it—and by ' we ' I mean university staffs, teachers, educational adinistrators and Ministry of Education officials. Few of us really believe in education as an academic subject, let alone education as a field for research. We have, in fact, got as much research as we deserve.

A number of things might be done to attempt to remedy this state of affairs, and among the most pressing—and the most promising—are the following:

(*a*) A change in Ph.D. regulations to permit teachers to work as part-time research students.

(*b*) The institution of more chairs in education in as many British universities as possible, the majority being in educational psychology.

(*c*) The provision of adequate aids to research in university departments (e.g., electric desk calculating machines, punched card equipment).

(*d*) The training of clerical staff, research assistants and teaching staff in the skills necessary for the effective use of electronic computers, and in particular tape-punching and programme-writing.

IV.—RESISTANCES, ORGANISATION AND COSTS.*

BY W. D. WALL
(*National Foundation for Educational Research*)

I.—THE SPEAR OF ITHURIEL.

IN his paper in this issue, Wiseman suggests that some at least of the fault is not in our stars but in ourselves that we do not command the resources necessary for educational research; those concerned with education do not demand from their universities, from the Ministry of Education and, indeed, from the community at large the resources necessary to develop research in education.

This is at least superficially true. But as one who has been concerned with educational research in a variety of different administrative and academic contexts, it seems to me that there is more to it than this. Education is almost the last major field of human endeavour where it is safe for the amateur to pontificate because everyone is possibly (or probably) wrong, even if every one is probably (possibly) right. One does not need to be expert in any real sense to have an opinion which will be listened to. It is, too, the field of human affairs in which those responsible for decision making differ most markedly from those for whose benefit the decisions are made and where still a very high proportion of administrators, inspectors, professors and so on have never shared as children and adolescents the experience of being educated in common with 95 per cent. of their fellows; or even in their professional career tried to teach them. In every part of our system from the classroom to the Advisory Committee, opinion, based upon more or less relevant, more or less objectively evaluated personal experience and on more or less genuine knowledge is the principal ingredient in decision making—and even in the choice of questions to which answers shall be sought.

Further, in most practical educational situations—from the choice of method whereby to teach reading to the determination of the proportion of an age group who shall be offered higher education—the taking of any decision is likely to be governed by at least two quite different elements—a value judgment which concerns the aims of education and a variety of more or less factual and practical considerations. These two elements are rarely fully disentangled and looked at separately and even more rarely is more than a small and often irrelevant proportion of the necessary facts available to the decision makers.

Excellent examples of what is meant are provided by current or immediately past controversies over such matters as comprehensive *versus* selective secondary education, allocation at 11 plus, or streaming. All these have been decided—and differently decided—with great conviction, basically on value grounds supported by a sufficient selection from the meagre facts and by opinions derived from experience. The experienced headmaster, Sir Cyril Norwood, finding his three types of pupil at 11; the groups who believe that the common school will abolish privilege, secure common humanity and tend to level up differences in ability: the teachers who believe that the needs of bright and not-so-bright children are best served by separating them into more or less homogeneous groups—all these and their opponents can adduce powerful arguments from experience backed up by estimations of what is a 'good' educational aim or philosophy, just, it might be remembered, as the 18th and 19th century educators could adduce opinion and experience in support of the notion of faculties, and of the value of Latin as a mental discipline.

---

* This paper has greatly benefited from the comments of Mr. H. L. Elvin, Director of the Institute of Education, University of London. The writer, however, accepts full responsibility for the views expressed.

Only rarely do teachers, educational administrators or advisory councils sit down and ask the three distinct, though related, questions—*What* should be aimed at ? (Values) ; is what we wish to have *attainable* compatible with other ' goods ' ? and *how* may the end be attained ? Rarely, of course, can these questions be asked or answered in their simple form but they do enable us to see that the realm of conviction, political, religious or social philosophy is essentially covered in the first question ; the other two can really only be answered empirically and it is here that research most clearly comes in, and is, at present, most clearly neglected. There is a fourth question to which an answer should be sought once a decision has been taken—are we achieving by the means chosen, the ends we proposed and are there any ' side effects,' good or bad ? Hitherto, this fourth question has tended to be answered by those responsible for taking the decision in the first place ; they go one better than cunning old Fury ; not only are they judge and jury, they are prosecution and defence, and the court is surprisingly lenient !

I am suggesting, therefore, that to some extent and probably unconsciously, there are defence mechanisms at work of a somewhat complicated kind. We do not relish the touch of Ithuriel's spear. Moreover, much of the resistance to the scientific study of education comes from those in university departments and in teaching who are professionally concerned with the improvement of the schools. Some experience a deeply emotional revulsion from any attempt to measure or predict human behaviour, to understand and, therefore, perhaps to control. Such opposition makes itself felt against all the social sciences more or less indifferently. It often expresses itself in moral terms—implying that it is sinful or dangerous to interfere with human beings, rejecting ' manipulation ' as immoral. More frequently these days it takes the form of a profound defensive belief masquerading as scepticism, which perhaps conceals a wish or a fear . . . " You can't change human nature," " You can't measure any of the important things," and so on. The difficulty with such widespread attitudes is that there is sufficient truth in them to obscure the main issues. It *would* be morally wrong to manipulate any human being for ends other than those which are ' good ' for him. However, all education is, in fact, manipulation ; its aim is to change, to structure thought and behaviour. Greater knowledge of human growth and learning, the refinement of techniques of influencing children will, we hope, merely make the process more efficient : in themselves, they will be neither more ' moral ' nor more ' immoral ' than any other scientific knowledge.

Those who express scepticism as to the possibility of research are more difficult to answer since only the success or failure of research work can, in fact, demonstrate their rightness or wrongness. The danger of much expressed scepticism about research in the disciplines that contribute to educational thought and practices is not, in fact, that it is scepticism—which is healthy—but that it is blind belief, a wish not to have one's ideas and practices disturbed masquerading as a commonsensical agnosticism. We know from the story of early renaissance science how such belief can inhibit enquiry and the progress of knowledge. History is strewn with the wreckage of ' facts ' hallowed by centuries of belief and even the infant science of psychology has been remarkably iconoclastic, in the last half-century.

More subtle still and more difficult to combat is the attitude which says—with truth—"research takes a long time, and anyway when we get the results the circumstances may have changed. Decisions have to be taken now ; we can't wait." Many who say this are genuinely friendly to an objective and empirical approach to their problems and painfully aware of the frequency with which they have to take action on the basis of inadequate knowledge. What they are liable to forget is that, while we may have to and, indeed, should act now in the absence of knowledge, this is neither an excuse for not attempting objectively and rigorously to measure and evaluate the results of our action nor a reason why we should not begin now to find somewhat more of the facts upon which our next decision should be based.

This summary and possibly otiose discussion perhaps enables us, while agreeing with Wiseman that in the present state of opinion and with university education departments staffed and organized as they are, no rapid expansion is possible, to see some of the deeper reasons and resistances which have to be overcome in the thinking of all those professionally concerned with education. It should also permit us—in the light of Peel's remarks on priorities and the experience described by Burroughs—to define some of the kinds of research and enquiry that are necessary, to serialise the organizational and conceptual problems involved and to suggest the financial and administrative structures which might be created effectively to support the kind of massive advance in knowledge and practice that could transform education over the next half century or so.

II.—EDUCATIONAL RESEARCH, STUDIES AND INVESTIGATIONS AND RESEARCH IN THE DISCIPLINES CONTRIBUTORY TO EDUCATION.

It seems to the present writer that we might begin usefully by distinguishing between educational research and research in those sciences which have a contribution to make to education. Leaving aside for the purposes of this discussion, research of a somewhat different kind in the history and philosophy of education and conventional comparative education, we can describe educational research, studies or investigation as being the empirical and applied study of things as they are, of alternative methods or organizations actually in operation, and the attempt to understand, evaluate and eventually to control an ongoing series of processes. We can also include in this, experimental education—that is the deliberate try-out of new methods, organizations and so on proposed on the basis of experience or of the results of research in other sciences.

This is *applied* research in the relatively classical sense of the term ; and it will take two broadly distinguishable forms : *operational* research (the study of what is) and *action* research (the setting up of experimental situations in the practical context of the educational system in order to test them.) Research of either type will call upon the concepts and techniques of a variety of sciences which, existing in their own right and with their own problems, are none-the-less in Claparede's sense, the sciences of education—general psychology, child development, psychometrics, sociology, statistics, economics and probably others.*

---

* Which others is partly a question concerning the development of the sciences themselves and partly a matter of the problem in hand. For example, scientific linguistics and semantics are clearly relevant to problems concerned with the teaching of reading, human genetics to the study of the relevance of particular kinds of educational method to children who may have qualitative differences in intellectual functioning. Similarly, philosophy helps in the precision of the aims to be achieved and, therefore, in the setting up of valid as well as testable hypotheses.

Research in the sciences contributory to education—much of it of the kind described by Peel—is not, perhaps, to be called ' pure ' in the sense used in the simpler sciences. From the present point of view it may be distinguished by not being concerned principally with an immediate, complex, practical problem but as being ' Fundamental ' in the sense that it is directed to the development and testing of a coherent body of theory in its own field, and to the fashioning of the appropriate instruments of study and measurement.

Within either of these two—educational research and research in the sciences of education—it is suggested that we may have ' hard ' research in Wiseman's terms or the sort of study which can to some degree at least be carried out by those experienced in education, but not specifically trained in the research methods appropriate to more elaborate enquiries. This latter type of activity has been likened by Hearnshaw to ' bird watching '—a kind of observational descriptive activity of a semi-scientific sort.

To the present writer, this latter seems a false, misleading distinction.

III.—RIGOUR.

The characteristic of research activity of a scientific sort lies, not in the *method* employed, but in the rigour with which any bias in the observer or inherent in the situation is excluded or allowed for, in the scrupulous testing of alternative hypotheses, in parsimony of interpretation and in the attempt to estimate margins of error in the measurements, estimates or observations on which it is based. How far any given enquiry can be rigorous and objective depends upon two things—the existence, accuracy and comprehensiveness of measuring and investigatory instruments ; and the degree to which those who plan the enquiry possess the conceptual framework and the knowledge of likely sources of error to analyse adequately the problem to be attacked ; and of these, the second is clearly more important than the first ; it amounts to an attitude of mind—rigorous but impartial agnosticism we might call it as against the prejudice (or scepticism) inherent in belief.

At this point it should be said most emphatically that in education, as in other social fields, many of the most important problems can at present only be investigated by dispassionate observation, with specific measurement techniques being used where appropriate and by weighting other factors in accordance with experience, judgment and insight. Such studies—if they are rigorously conceived and especially if they are conducted within a strong and sound conceptual framework, have the basic characteristic of scientific research. We might go further and suggest that they are more ' scientific ' than those which attempt to leave out of the study any aspects which are too subtle and complex for measurement (at present at least) or than those which use falsely precise measures.

In the early stages of any science the concepts, the investigory techniques and the measures are lacking. 'Bird watching' is an essential preliminary to more elaborate studies ; but it need be no less scientific—i.e., rigorous and objective—for that. Its margins of error will be larger, its predictions less sure and the conceptual framework which it begins to build more vague and fluid at first than one would hope to attain by the successive approximations to the truth which is the march of any developing science. Some of the difficulty of the social scientist and particularly of those who work in the field of educational research, is that, knowing our techniques to be much less powerful and objective than we would like, too many of us seek to anticipate criticism by adopting not necessarily appropriate methods from the natural and 'exact' sciences,* by limiting our enquiries to those trivialities which lend themselves to a more-or-less artificial and spurious objectivity of examination, or by confining ourselves to the role of critic giving no hostages to fortune or our colleagues.

It is quite true that almost any real issue in education or in child development presents a dangerous and threatening situation; the problems to be studied cannot be much simplified without radically changing them: the variables concerned are rarely independent, rarely simple and practically never can they be systematically eliminated or varied. The best measures we have—for example, a test of a simple educational attainment—are remarkably blunt and imprecise instruments. Few of us, indeed, sufficiently dominate the various related and often incipient bodies of knowledge and experience necessary both to formulate adequate hyppotheses and to estimate the sources of error. This makes it all the more necessary that strong, well-staffed departments of research and higher degrees should exist which can provide a coherent programme of researches and studies within the framework of which the individual M.A. or Ph.D. student can investigate a real and complex problem rather than, as too often at present, playing safe by choosing an innocuous more or less trivial topic or a well-trodden path.

IV.—RESEARCH CHAIRS? RESEARCH DEPARTMENTS?

Wiseman suggests as some part of the remedy for this the establishment of research chairs in the universities. The idea is attractive, even seductive. There is, however, room for serious doubt as to its effectiveness as the principal means of developing research. There are, as has been suggested earlier, a number of sciences which have contributions to make to education and it is doubtful whether one research chair per department would be anything like adequate. In spite of the invidious choices which would be involved, it would probably be better to develop strong research departments adequately staffed and financed, in a few universities, rather than spread research chairs and resources out too thinly. A degree of specialisation might then be developed and M.A. and Ph.D. students be encouraged to go to a department whose ongoing programme of research is in line with their interests. Research of any complexity and continuity depends not merely upon the creative research worker, or on the stable team concerned with a broad area of study; nor is the 'ironmongery' of calculating machines and electronic computers the whole answer. Any worthwhile research in education involves an immense amount of clerical and administrative work. Experience is beginning to suggest that each research worker requires at least two experienced clerical assistants and a small team of computers.

University administrators have not yet accustomed themselves to the idea that it is uneconomic to oblige a relatively highly paid member of the professional staff to do work which can be done at less cost (and more efficiently) by a clerk or secretary. The dangers in our present habits of thought about research are that because we do not grasp the nature and the magnitude of the problems before us, we shall plan, staff and finance research on a scale which will condemn us either to confine our studies to probes of a pilot kind or to limit them to largely survey kinds of studies which, though interesting and valuable, fall short of what is necessary for the development of comprehensive theory.

---

* Whose 'exactness' is merely relative and by no means absolute.

V.—SOME CURRENT RESEARCH AND ITS LESSONS.

We may, perhaps, illustrate this by considering some kinds of educational research which are at present going on. The National Foundation for Educational Research has undertaken two National surveys of educational attainment, one in 1955 and the other in 1960; further similar surveys are planned at quinquennial intervals. The rationale and methods of these surveys have been described by Pidgeon in this *Journal* and elsewhere. The value of the results which are being obtained is unquestioned and they supplement both in breadth and in depth the pioneer surveys of Godfrey Thomson and the Scottish Council for Research in Education. Such surveys are based upon group tests administered for the most part by teachers; and the scores are related to a considerable array of other variables, such as socio-economic indices, parental occupation, types of school organization, size of class and, in the case of Wales, linguistic background. Information on most of these variables has to be obtained by questionnaires and inevitably is gathered in a simplified form. The analysis of such information can result in the discovery and establishment of suggestive associations; but it can tell us little of dynamic interactions and little or nothing about causes. From such data predictions about the expected behaviour of groups can be made with known likelihoods of error; little light, however, can be thrown upon individual variation or the reasons why the exceptions occur. For this, two further steps are clearly necessary; the one would involve continued close skilled study of individual children; the other would imply the systematic experimental variation of a major variable—for example, type of school organisation. Immediately such a conception is accepted, it becomes clear that the staffing involved is both more considerable in number and different in quality from that required for a survey—and the cost very much higher. Whereas a major survey costs about £4—£5,000 per annum its enrichment in depth would probably double the expenditure.

Similar remarks may be made about operational research of the type undertaken by the N.F.E.R. in Reading (Dr. J. M. Morris) and Arithmetic/Mathematics (Dr. J. B. Biggs). In the first case sizeable groups of children have been studied individually, followed up over a period with marked enrichment of the results, but it has not so far been possible to develop a complementary action type of research which, in the light of knowledge gained from the broad surveys, would vary teaching method systematically over a number of years. In the case of the arithmetic/mathematics enquiry, a complementary action research has been started, but because only one fully trained worker can be employed on it, the enquiry will inevitably lack the depth which would be given—for example—by prolonged classroom observation or the individual testing of a sample of children at intervals.

Significant advances in our knowledge of education are likely to come when we can combine some or all of the following in our research designs:

(*a*) Broad background sampling surveys.

(*b*) Intensive case studies of individual children in the light of such surveys.

(*c*) Experimental situations deliberately set up—e.g., teaching method experiments.

(*d*) The follow-up of groups over a considerable period (5-10 years).

Furthermore, most educational problems involve for their study and solution at least the sciences of experimental pedagogy,† educational psychology, child development, sociology, and probably economics; because of this other things of a structural and administrative kind are involved. The research must be planned and probably conducted by a team representing most of the relevant disciplines—even if some are rudimentary and even if leadership lies with one individual. It must be affiliated to an institution which can ensure continuity and follow-up so that the departure of a research worker or professor does not lead to the abandonment of the research. And it must be supported by an efficient, continuing and (relatively to current university standards) lavish administrative and clerical staff. If to this we add the point made by Wiseman that machines—desk calculators and electronic computors—are also essential, not on the basis of temporary loans but as an integral part of the research set-up, we begin to see that considerable resources are involved in any expansion.

The current danger is that we shall be timid and—because we do not immediately have a source of highly trained research workers—expansion will take the form of too little and too late. This—as we can see by a study of the developments of the last ten years—leads to a hotch-potch development. Some are tempted to go for the more-or-less uncontrolled study or experiment which can be quickly conducted and written up rather than take the risk of embarking on a long-term and reasonably rigorous research.

Here, perhaps, we should pause and count the likely costs. An adequate method study conducted in primary schools and supported by the full voluntary co-operation of teachers and L.E.As. but not containing any element of individual psychological study of children (which would be costly in time) costs at a minimum between £5,000 and £6,000 a year in salaries, administrative costs, materials and so on. Field work could not take less than two to three years; this with a preparatory year and a year or more for the analysis and writing up of the results, means that the lowest real cost will be £25—£35,000. If, however, such a project is enriched by greater complexity in the variables studied—for example, an analysis of children's motivations, interests, attitudes, etc., as well as their attainments—and by the study of the effects of a method on individual thought processes; and if provision be made for a follow-up of the subjects over a further three or five years, then the costs would be at least doubled.

Costs of the order of £10—£12,000 a year for a project only seem high because in educational research we have never had resources bearing any relation to the possibilities of useful research let alone to needs. Industry is prepared to spend more than this in investigating a likely market for one product. Almost all the valuable work going on at present is underfinanced. As far as I know, no research unit concerned with educational research possesses even the simplest machinery for card-sorting: and many workers consider themselves lucky to have a hand calculating machine. Apart from small specialised units in four or five universities, the N.F.E.R. is the only organisation which can ensure continuity in a long-term research programme. Can we truly say, in the face of our increasing rate of educational expenditure, that—say—a hundred projects of the kind described in the previous paragraph are more than could be mounted successfully over the next five years? or than are urgently necessary?

---

† I apologise for this continental phrase, but we do not seem to have a term for the devising and testing of specific methods, materials and so on, which is an as yet relatively undeveloped aspect of educational study in this country.

VI.—NATIONAL ORGANISATION OF RESEARCH.

It is instructive to consider what such an effort might mean financially, organisationally and in terms of our attitudes. If the universities are to play a significant part in the development of educational research, then certain departments and institutes will have to have fully staffed and equipped departments of research and higher degrees. Most of their staff will have to be full-time research workers with only occasional teaching duties of the familiar kind—lecture courses, seminars and teaching practice supervision—though one would hope that colleagues concerned with teaching would also participate in research. Such research departments will have a double function—that of conducting research, on the necessary scale and over considerable periods of time, devoted to a broad field of study, and that of providing a thorough basic training for research workers. A student wishing to read for an M.Ed. or Ph.D. would, in fact, have to choose his university in function of the research interests of the department of education concerned rather than join a department and do research of his own choice irrespective of the competence or interest of his supervisors.

While it is clear the principal discipline involved in most empirical research in education is educational psychology, the staffing of the university research units—and the provisions of senior posts, readerships and chairs—will have to reflect the fact that there is a range of sciences which should contribute to educational study and research. We should expect to find sociologists, statisticians, economists, social psychologists, pure psychologists and others in the teams, differently combined according to the area of research. We may have to go much further than this. Much research basic to education should, in fact, be inter-departmental, involving colleagues in other social sciences, in medicine, physiology and so on. The existence of strong departments in, say, the basic sciences of psychology, sociology and statistics might—as well as the strength of the department or institute of education itself—be a criterion determining the placement of a given research unit.

It is unlikely that we could finance, equip and staff at every university a worthwhile department of research and higher degrees on this scale ; and it is better to have fewer good ones, than to disperse resources, in the interests of ' fairness.' Moreover, there are other and less expensive forms of research activity in education which might appropriately be placed within smaller departments.*

Much of the research to be carried on should be of the kind described by Peel as ' probes ' ; much of it one would expect to be concerned with the development and testing of theory · some of it would be basic in the sense that it concerned itself with the shaping of instruments and techniques ; and some of it would be applied in the broad sense that it would experiment in considerable numbers of schools or with large samples of children.

---

* Education is a field of applied study to which research and study of various kinds—not all of it based on measurement, experiment or objective observation—has a major contribution to make ; but this contribution is not unique. What we call the theory of Education will always consist of our reflective and critical thinking about education in this light of theory and research in the basic disciplines that contribute to educational study. Ultimately, even the smallest educational decision in practice will be based upon an assessment of facts, experience, research results and so on in the light of a system of values, a philosophy of education and of society. Hence, every education department might be expected to be a centre of development of educational theory in these senses.

However, university research is not likely to meet all needs or satisfy all demands. Much research must be of an operational or action type—concerned with the working of what is, or, experimenting, with novel techniques, organisations or methods—and must be national in its sampling as well as in the co-operation it evokes. For this, institutions much more fully staffed and equipped and similar in standing to the present National Foundation for Educational Research or the Scottish Council for Research in Education are clearly necessary. Each of these institutions represents a broad partnership between educational authorities, teachers, universities and the Ministry of Education and each is well placed to do work on a national scale, to provide services of co-ordination, and to act as a clearing house for the exchange of information between research workers and for the rapid feedback into the educational system of knowledge gained from research. In the light of needs, neither the S.C.R.E. nor the N.F.E.R. is staffed or financed to fulfil all the tasks they might with advantage do; but each has the nucleus necessary for expansion, and, what is perhaps even more important, the know-how and the goodwill of the schools. Such national institutions might be concerned not only with research of the type they now do, but with such matters as the testing and development of new educational material, e.g., teaching machines, educational television, textbooks—controlled experiments with new curricula, as well as with assistance in research, consultative and information services for teachers, L.E.A's and the Ministry of Education. They should also play a much more considerable part than they now do in the training of research workers.

The third leg of an adequate organisation for educational research is one frequently overlooked. Local education authorities are, much more frequently than ever before, involved in taking crucial educational decisions which involve considerable expenditure and affect the happiness and success of large numbers of children. Even where some of the evidence exists to provide the background for such decisions, local circumstances are likely heavily to modify it. Similarly, decisions made and implemented should be subject to something more than casual impressionistic evaluation. Finally, there is at present a great deal of information gathered in L.E.A. offices—about children, about teachers, about expenditure, about the results of examinations, for example—which, once it has met its *ad hoc* use, drains into the sand. All this seems to point to the need for L.E.A's. of any size to have small research units of their own—probably as a development of their psychological services which have the skills to undertake the work, the close contact with the schools which makes the gathering of data and the dissemination of information relatively easy and a certain measure of independence.

The kind of organisation outlined here will not come about spontaneously. Aware of the needs and of the history of the development of research in Medicine and Agriculture, the Parliamentary and Scientific Committee recently put forward a memorandum to the Ministry of Education and Minister for Science. The suggestion there was made that the Government should set up a Council for Educational Research responsible directly, as are the M.R.C., the A.R.C. and the D.S.I.R. to the Privy Council through the Minister for Science. On such a Council the Ministry of Education, Local Education Authorities, Teachers, Universities and others would be represented by presumably expert and knowledgeable people who could undertake the task of planning development. The then Minister of Education, Sir David Eccles, did not think well of the idea or thought it inopportune. He had already set up within the Ministry a ***Research and Intelligence Branch*** endowed with a small but—one would like to hope—

growing budget—thus repeating the experience of Michael Sadler and the early days of the A.R.C. Sadler's unit closed down and the A.R.C. had to be—with some difficulty—detached from the Ministry of Agriculture before it could fully develop. The work of the Ministry's R. and I. Branch so far, in spite of its limited resources, has been excellent; and probably the Ministry will always need to have such a section. However, one must beg leave to ask whether this is enough and whether the development of research in education is best served by a small section in any Ministry.

VII.—CONCLUSION.

In different ways the contributors to the present symposium have stressed that our present research effort in education is pitifully small and that what there is of it is patchy. They have stressed that, in education, there are at least three functions which research should serve and which are not necessarily always separate: that of advancing theory; that of testing the results of decisions and modifying practice while it is in course; and that of improving the professional skills of teachers and others through participation in research. We are not perhaps so well agreed about how far and how fast we could go, if the financial resources were available. We all point to the consequences of the neglect of the past decades—the absence of trained staff, of the organisational structures and even the concepts, and of the expensive machinery for statistical analysis. The present writer would agree that, if we were given a million pounds tomorrow, there would be serious embarrassment in attempting to use it wisely on research. This, however, should not be allowed to become an excuse for not planning now for a steady and fairly rapid build-up of effort over the next five or ten years; nor should we encourage the belief that more and more uncontrolled experiments and well-intentioned studies of this or that without any attempt at rigorous control of the results are, in the absence of fully developed techniques of measurement and evaluation and of highly trained workers, a satisfactory substitute.

THE RELATIONSHIP BETWEEN INHIBITION OF MOTOR RESPONSE AND COGNITIVE PERFORMANCE*

BY C. C. ANDERSON

(*University of Alberta*)

SUMMARY. Tests of motor inhibition, cognition, socio-economic status, psychological control and dogmatism were administered to a sample of grade 8 students (average age 14 years) so that two hypotheses could be tested : firstly, that the capacity for motor inhibition was correlated with level of cognitive performance, and secondly, that such a correlation could be explained in terms of child rearing practices characteristic of parents of relatively superior socio-economic status, which involve the positive reinforcement of motor inhibition, the presence of positive reinforcement being signified by a negative correlation between inhibition ability and dogmatism. The first hypothesis was supported, upper inhibition groups showing a consistent and occasionally significant superiority in cognitive performance over lower inhibition groups. The second hypothesis was not supported, the correlation between dogmatism and control retaining its significance and even its magnitude, despite the fact that the influence of socio-economic status had been partialled out. An explanation was attempted of the significant correlation between inhibition capacity and score on dogmatism.

I.—INTRODUCTION.

SINGER (1955) has impressively documented the fact that many psychologists notably Freud (1950, p. 553, originally 1900 ; and 1959, p. 16, originally 1911), Thurstone (1924) and Harlow (1958, p. 282) have hypothesized an association between the capacity to inhibit motor responses and level of cognitive performance. Evidence for this hypothesis has been provided by Meltzoff and Levine (1954), Levine, Glass and Meltzoff (1957), Levine, Spivack and Wight (1959) and Spivack, Levine and Sprigle (1959). The most obvious hypothetical antecedent condition of this relationship is, as Singer (1955, pp. 263-264) suggests, a pattern of child-rearing in which the parents have reinforced any attempt on the part of the child to delay the gratification of fundamental drives and to inhibit associated motor activity. The relevance of these investigations to education has not been noted partly because the subjects studied have been adults and partly because experimental and research tests of cognitive performance have been used in place of the more conventional intelligence tests. The present study is an attempt to remedy this situation by validating these two hypotheses as manifested in the responses of adolescents and pre-adolescents to some of the tests of motor inhibition used in the previous studies, and to cognitive tests of the more usual standardized sort.

The first hypothesis, that there is a relationship between the capacity to inhibit motor response and level of cognitive performance, was tested directly by comparing the extent of cognitive performance of a group of marked inhibitors with that of a group of marked non-inhibitors. The second hypothesis that the foregoing relationship is a product of a child-rearing pattern in which delay of motor responses, especially those associated with the gratification of fundamental drives, is reinforced and the opposite behaviour either punished or ignored, was tested indirectly in the following way. Davis (1952) has concluded

---

* The writer is indebted to the University of Alberta Research Grants Committee for financial aid for this research.

that this capacity to inhibit motor response is a product of parental punishment of undesirable behaviour, a practice which is likely to activate in the child a state of ' socialized anxiety ' which motivates him to behave in parentally-approved ways. If Davis is correct in his contention that such an upbringing is characteristic of urban middle-class parents and if Singer's hypothesis about the child-rearing antecedents of the relationship between motor inhibition and cognitive performance is also correct, then the group of motor inhibitors must be significantly superior to the group of non-inhibitors in socio-economic status as well as in cognitive performance.

However, this prediction can also be made without using Davis's concept of ' socialized anxiety ' and by stressing the positive or Skinnerian approach to reinforcement. For example, Maccoby, Gibbs, *et al.* (1954) conclude, from their study of methods of child-rearing by mothers in two social classes (upper-middle and upper-lower) in Boston, that upper-lower mothers are more likely to use negative reinforcement techniques in disciplining the child (physical punishment, ridicule, deprivation of privileges), the upper-middle mothers being far more permissive and imposing fewer restrictions. If this is the case and if motor inhibition can be learned as a result of positive parental reinforcement, it might again by hypothesized that the inhibitors would be markedly superior to the non-inhibitors on a measure of socio-economic status. Therefore, to facilitate a less ambiguous explanation of the child-rearing conditions underlying any observed relationship between inhibition and socio-economic status, a test of dogmatism was also administered to the sample of subjects. This was a move dictated by the following consideration. Frenkel-Brunswik's (1954, pp. 236-237) description of the upbringing of authoritarians stresses the presence of negative reinforcement and is remarkably similar to Davis's picture of middle-class child-training in terms of socialised anxiety. If, therefore, the present research establishes positive relationships between inhibition and socio-economic status and between inhibition and dogmatism, then Davis's position will be supported. If, however, Maccoby is correct in her description of the relatively permissive middle-class child-rearing practices which reduce the child's ambivalence towards the parents, the hallmark of the authoritarian (Sanford, 1956, p. 278), then the relationship between inhibition and dogmatism will be negative, a point already made plausible by Gough (1956, p. 192).

Accordingly, the two hypotheses tested in the present study are : firstly, that motor inhibitors are significantly superior to non-inhibitors in cognitive performance ; secondly, that they are also significantly superior in socio-economic status, but significantly inferior in dogmatism, the latter two predictions encompassing the second hypothesis testing the validity of the Maccoby version of child-rearing practices by parents of higher socio-economic status.

II.—METHOD OF INVESTIGATION AND TESTS USED.

The testing was carried out in three large schools which had been chosen in 1960 by a graduate student, Dr. W. B. Elley,* as typical of the grade VII (average age 13 years) school population of Edmonton, Alberta. Elley selected this sample after demonstrating that it did not differ significantly in age, intelligence, occupational level and sex from a random sample of 432 pupils tests, took the following tests in 1960 and the scores were made available to the present writer one year later by Dr. Elley: Raven's (1954) Standard Progressive Matrices; the Lorge Thorndike Non-Verbal Intelligence Test, level 4 (1957); the California Test of Mental Maturity, Short Form, Elementary (1957); and a Home Index questionnaire, adapted from one devised by Gough (1946). In 1961, when the pupils were in grade VIII, the present writer administered the following tests: The motor inhibition (MH) tests were borrowed with slight changes from Meltzoff and Levine and from Neel (1960). In MH1 and MH3, the pupils read a paragraph describing an exciting episode (a motor race in MH1; a lion hunt in MH3) and were then asked to copy out the paragraph as slowly as possible, taking care to keep the pencil always moving on the paper. The instructions were repeated twice, and the time allotted for slow writing on each test was five minutes. In the second test (MH2), the pupils were encouraged to write as quickly as possible the sentence " Edmonton is in Canada " for two minutes, and then were switched to writing the same sentence on another page as slowly as possible for two minutes. The first, second (MH2A) and third tests were scored by subtracting from 100 the total number of letters written by the pupil. The second test was also scored (MH2B) by multiplying by 10 the ratio of the number of letters written in the fast-writing phase to the number written in the slow-writing phase. Two other tests were administered, a Psychological Control Questionnaire (PCQ), previously used by Luft (1959) and Shippee-Blum (1959) and designed to measure the capacity to control impulse gratification, and an abbreviated version of Rokeach's (1960) test of dogmatism devised to suit the pupil's level of maturity

---

* The present writer is indebted to Dr. Elley for his provision of the test scores. Further details of the sampling can be found in Elley's thesis, " A Comparative Study of Socio-Economic Bias in Selected Intelligence Tests."

III.—RESULTS.

The distributions of scores for each of these inhibition indices (1=MH1, 2=MH2A and 3=MH3) were normalised and the intercorrelations ran: $r_{12}=\cdot834$; $r_{13}=\cdot902$ and $r_{23}=\cdot879$. The three tests were regarded as reasonably reliable and suitable for research purposes.
drawn from the city-wide grade VI testing carried out in 1959. The present sample, consisting of 258 of Elley's 271 students who had completed all the

Ten scores were available for each student, one each on MH1 and MH3, two on MH2, and one each on the following variables: Home Index (HI), Raven, Lorge-Thorndike (L-Th.), California Verbal (Calif.), Psychological Control Questionnaire (PCQ) and Dogmatism (Dogm). The top and bottom 15 per cent. (N=38) of cases in the four inhibition distributions and on a composite distribution based on the total of the four scores were isolated into upper and lower inhibition groups, and the significance of the differences between their average performances on the other six variables was tested by analysis of variance, the relevant F ratio appearing in Table 1.

Certain conclusions may be drawn from these results. The upper inhibition groups perform consistently better on the cognitive variables than the lower inhibition groups, especially the groups isolated on the basis of performance on MH2A, and application of the sign test (fourteen of the fifteen F ratios are in the expected direction) permits the conclusion that there is a significant difference between the cognitive performance of groups differentiated on the basis of measures of motor inhibition. The first hypothesis is validated. Probably all F ratios would have been significant and recourse to the sign test unnecessary if all inhibition measures had been obtained like those in MH2 after a significant amount of arousal (Hebb, 1958, p. 250) had been activated. Lindsley (1957, p. 80) has argued in the same way that the insignificant relationship established

TABLE 1

F RATIO ESTIMATES OF THE SIGNIFICANCE OF THE DIFFERENCES BETWEEN UPPER AND LOWER INHIBITION GROUPS ON SOME DISCRIMINANT VARIABLES.

	HI	Raven	L-Th.	Calif.	PCQ	Dogm.
HM1	·052	·142	·938	·455†	·868†	6·195*†
MH2A	8·59**	2·91	8·31**	7·87**	2·58	5·07*†
MH2B	·83	·279	2·27	1·14	2·50	7·33**†
MH3	1·77	·473	3·08	1·67	·001	10·75**†
MH Composite	11·97**	·478	1·80	2·35	·100	12·528**†

** Significant at the 1 per cent. level.
* Significant at the 5 per cent. level.
† Lower inhibition group mean greater than upper inhibition group mean.

between traditional EEG measures and intelligence would not appear if the former measures were made after arousal had been activated. It is reasonable to assume that a measure of inhibition is most valid when taken during a time of enforced inhibition.

A second conclusion is that the upper inhibition groups are significantly less dogmatic than the lower inhibition groups, and also tend to show higher scores on the PCQ. Further, the upper groups isolated on the basis of MH2A and the composite index are significantly superior to the lower groups in socio-economic status, and the groups derived from the other three indices show a consistent trend in this direction. However, the second hypothesis is not given much support. The association between inhibition and dogmatism is clearly stronger than that between inhibition and socio-economic status because significant differences between the two groups in dogmatism are observed on three occasions (MH1, MH2B, and MH3) while there are no corresponding differences in socio-economic status. It was, therefore, decided to test the validity of the second hypothesis by calculating, for MH2A and composite scores, where the association between inhibition and socio-economic status is strongest, the extent of the correlation between inhibition and dogmatism with the influence of socio-economic status partialled out. Clearly, if the significant negative correlation between inhibition and dogmatism is a product largely of the child-rearing practices associated with socio-economic status, then this partial correlation will be insignificant.

TABLE 2

PARTIAL CORRELATIONS DERIVED FROM TWO SETS OF DATA BETWEEN INHIBITION (MH) AND DOGMATISM WITH SOCIO-ECONOMIC STATUS PARTIALLED OUT.

	MH2A	MH Composite
^{r}MH–Dogm.	—·424	—·536**
^{r}MH–HI	·443**	·345**
^{r}Dogm.–HI	—·148	—·147
Partial ^{r}MH–Dogm.	—·403**	—·522**

We may conclude from these findings that the association between motor inhibition and dogmatism is independent of socio-economic status and that, therefore, the second hypothesis is not supported.

It might be argued that this significant partial correlation is largely a product of sex differences in the capacity for motor inhibition. There is evidence (Carrigan, 1960, p.334) for the widely-accepted view that females are significantly antly more impulsive than males; possibly, therefore, there are significantly more females in the lower inhibition groups and they may be more dogmatic than males. Table 3 provides evidence bearing on the first part of this contention.

TABLE 3

THE NUMBER OF MALES AND FEMALES IN ALL FOUR UPPER AND LOWER INHIBITION GROUPS.

		Sex		Totals
		F	M	
Inhibition Groups	Upper	82	108	190
	Lower	106	84	190
	Total	188	192	380

The chi-square of 6·064 is significant at the 5 per cent. level and it may be concluded that, for these data at least, females are markedly more impulsive than males. But are these impulsive women significantly more dogmatic than lower-group males, and, by implication, are the less impulsive males in the upper groups less dogmatic than upper-group females? The average scores on the dogmatism test of males and females in all five upper and lower inhibition groups were: upper females 37·5 (S.D. 13·4), upper males 35·8 (S.D. 14·5), lower females 45·4 (S.D. 13·6) and lower males 49 (S.D. 14·5). The within-inhibition groups mean differences are insignificant and, in any case, the sex means run in a direction opposite to that predicted by the argument under consideration. Clearly, there is no evidence for the contention that the females who preponderate in the lower-inhibition groups are significantly more dogmatic than their male counterparts.

A final point remains to be made. It might be argued that the relationship between inhibition and cognition is curvilinear and that a group chosen from the middle of the distribution of any inhibition scores would be either significantly superior or significantly inferior to the upper and lower groups in average cognitive performance. If this were the case, then analysing the performance of the two isolated groups *only* would merely obscure rather than clarify the relationship. However, the average cognitive performance (on Raven, Lorge-Thorndike and California Verbal) of five median inhibition groups consistently fell between the average of the corresponding five upper and lower inhibition groups, and it was concluded that the relationship between inhibition and cognition was linear rather than curvilinear.

IV.—DISCUSSION.

The relationships established between inhibition, cognition and dogmatism are congruent with a pattern of results from personality studies in which extraverts are designated non-inhibitors (Eysenck, 1957; Carrigan, 1960) relatively inferior to introverts in cognitive performance (Lynn, 1958) and significantly more authoritarian (Eysenck, 1954). Reinforcing this point, Messick and Hill (1959) have demonstrated that the impulsive person acts quickly without too much reflection, generalizes on the minimum of evidence and reaches perceptual closure quickly, characteristics which impair cognitive performance and which are aspects of the intolerance of ambiguity which Frenkel-Brunswik (1949) has claimed to be fundamental to authoritarianism. The present results also fit neatly into a neuropsychological model of motivation recently elaborated by Berlyne (1960), who contends, borrowing heavily from Hebb, that stimulation of the reticular arousal system (Lindsley, 1957 ; Magoun, 1958) facilitiates both motor and cortical activity, and that much of the organism's behaviour is designed homeostatically to restore the original and presumably optimal level of arousal. Part of this control or arousal-reduction behaviour consists of voluntary and involuntary inhibitory impulses sent by the cortex to the RAS to counteract the original arousal. Anxiety arises, according to Berlyne (1960, p. 181), if the cortex is caught momentarily off balance and cannot exercise its restraining influence immediately to restore equilibrium ; an extrapolation from this position might take the form that anxiety will also arise whenever the arousal system is so persistently and strongly activated that it taxes the individual's power of corticoreticular control. The arousal potential (*ibid.*, p. 179) most capable of stimulating such an unpleasantly high arousal level is that involving ambiguity or uncertainty (*ibid.*, p. 206) and consequently, the impulsive, easily-aroused, extraverted person, who is incapable of tolerating such a state of uncertainty to which he is liable to be subject, will tend to find arousal-equilibrium in dogmatism. He will also be characterized, if Berlyne's model is correct, by the relatively powerful motor impulses and impaired cognitive performance demonstrated in the present experiment. This is the same contrast as Couch and Keniston (1960) make between ' yeasayers ' and ' naysayers,' the main trait discriminants being degree of oral and motor speed and activity (1960, p. 163), ' impulse control ' (*ibid.*, p. 167) and ' intellectual orientation,' the extreme poles of the last discriminant being the " . . . yeasayer's preference for blunt, straightforward, and extreme opinions . . . " as against the naysayer's cautious and logical analysis of questions (*ibid.*, p. 168).

V.—REFERENCES.

BERLYNE, D. E. (1960). *Conflict, Arousal and Curiosity.* New York : McGraw-Hill.

CARRIGAN, P. M. (1960). Extraversion—introversion as a dimension of personality a reappraisal. *Psychol. Bull.*, 57, 329-360.

COUCH, A , and KENISTON, KENNETH (1960). Yeasayers and naysayers : agreeing response set as a personality variable *J. Abnorm. Soc., Psychol.*, 60, 157-174.

DAVIS, A., (1952). *Social Class Influences Upon Learning.* Cambridge : Harvard University Press.

EYSENCK, H. J (1954). *The Psychology of Politics.* London . Routledge and Kegan Paul.

EYSENCK, H J. (19 9). *The Dynamics of Anxiety and Hysteria.* London : Routledge and Kegan Paul.

FRENKEL-BRUNSWIK, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *J. Pers.*, 1949, 18, 108-143.

FRENKEL-BRUNSWIK, E. (1954). Further explorations by a contribution to ' The Authoritarian Personality? In Christie, R. and Jahoda, Marie (Edits.) *Studies in the Scope and Method of ' The Authoritarian Personality.'* Glencoe, Illinois : The Free Press

FREUD, S. (1900). *The Interpretations of Dreams* New York : Macmillan, 1950 edition.

FREUD, S. (1911). Formulations regarding the two principles of mental functioning. In *Collected Papers*, Vol. IV, pp. 13-21. New York : Basic Books, 1959.

Gough, H. G. (1949). A short social status inventory. *J. Educ. Psychol.*, 40, 52-56.

Gough, H. G. (1956). A scale for a personality dimension of socio-economic status. In Welsh, G. S., and Dahlstrom, W. G. (Edits.) *Readings on the MMPI in Psychology and Medicine*. Minneapolis . University of Minnesota Press.

Harlow, H. (1958). The evolution of learning. In Roe, A. and Simpson, G. (Edits.) *Behavior and Evolution*. New Haven : Yale University Press.

Hebb, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychol Rev.*, 62, 243-254.

Levine, M., Glass, H , and Meltzoff, J. (1957). The inhibition process, Rorschach human movement responses, and intelligence. *J. Consult. Psychol.*, 21, 41-45.

Levine, M., Spivack, G., and Wight, B. (1959). The inhibition process, Rorschach human movement responses and intelligence : some further data. *J. Consult. Psychol.*, 23, 306-312.

Lindsley, D. B. (1957). Psychophysiology and motivation. In Jones, M. R. (Edit.) *Nebraska Symposium on Motivation*. Lincoln, Nebraska : University of Nebraska Press

Luft, J. (1957) Psychological control patterns within families. *J. Consult. Psychol.*, 21, 206.

Lynn, R. (1958). Two personality characteristics related to academic achievement. *Brit. J. Educ. Psychol.*, 29, 213-216.

Maccoby, Eleanor, E., *et al.* (1954). Methods of child-rearing in two social classes. In Martin, W. E., and Standler, C. B. (Editors) *Readings in Child Development*. New York : Harcourt, Brace.

Magoun, H. W. (1958). *The Waking Brain*. Springfield, Illinois : Thomas

Meltzoff, J., and Levine, M (1954) The relationship between motor and cognitive inhibition. *J. Consult, Psychol.*, 18, 355-358.

Messick, S., and Hill, J. R. (959). Objective measures of personality ; cautiousness and intolerance of ambiguity. *Educational Testing Service Research Bulletin*, February, 1959.

Miller, D. R., and Swanson, G. E. (1958). *The Changing American Parent*. New York : Wiley.

Neel, Ann F. (1960). Inhibition and perception of movement on the Rorschach. *J. Consult. Psychol.*, 24, 226-230.

Rokeach, M. (1960). *The Open and Closed Mind*. New York : Basic Books.

Shippee-Blum, Eva M. (1959). The young rebel : self-regard and ego-ideal. *J. Consult. Psychol.*, 1959, 23-, 44-50.

Singer, J. L. (1955). Delayed gratification and ego development . implications for clinical and experimental research. *J. Constult. Psychol.*, 19, 259-266.

Spivack, G , Levine, M., and Sprigle, H. (1959). Intelligence test performance and the functions of the ego. *J' Constult. Psychol.*, 23, 428-431.

Thurstone, L. L (1924). *The Nature of Intelligence*. London : Kegan Paul, Trench, Trubner.

(*Manuscript received, January 10th*, 1962.)

THREE EXPERIMENTS USING THE JUNIOR MAUDSLEY PERSONALITY INVENTORY

I.—NEUROTICISM AND EXTRAVERSION IN SCHOOLBOYS AS MEASURED BY THE JUNIOR MAUDSLEY PERSONALITY INVENTORY

By M. PAULINE CALLARD and CHRISTINE L. GOODFELLOW
(University of Exeter)

Summary. The Junior Maudsley Personality Inventory was administered to 3,559 schoolboys between the ages of 11 years to 14 years 11 months, in secondary schools in Devon. The results obtained indicate that there is an association between achievement of status in the educational hierarchy and scoring on Neuroticism. Lower status groups (e g., secondary modern schools and less intelligent groups within all schools) were shown to have higher Neuroticism averages than the higher status groups (grammar schools and more intelligent groups within schools). The evidence was that at the top of the educational hierarchy there is a tendency for Neuroticism and Intelligence to be positively correlated for individuals whilst at the bottom of the hierarchy they are significantly negatively correlated. Neuroticism and Extraversion were both found to vary slightly with age. The younger age groups have higher Neuroticism averages but lower Extraversion averages.

I.—Introduction.

The Junior Maudsley Personality Inventory is designed by Messrs. Furneaux and Gibson (1961) to measure Neuroticism and Extraversion of children within the age range of about 9 to 14 years. It is based on Eysenck's theory of personality factors (Eysenck, 1953) and is the only personality test of this kind available for English children at present.*

The J.M.P.I. consists of fifty re-arranged and some re-phrased items of the Pintner Children's Personality Inventory (Pintner, *et al.*, 1938), making it less sensitive to response bias, directed only to the assessment of Neuroticism and Extraversion and more easily comprehensible to English school children. It was standardised by administering it to seventy-eight boys and seventy-eight girls between the ages of 9 and 14 years, under both individual and group conditions. The distribution of I.Qs. for the group was positively skewed. The resultant scores were—Extraversion: mean 12·39, S.D. 3·54; Neuroticism: mean 7·35, S.D. 3·54. Separate scores were not distinguished for boys and girls.

The aim of this research was to establish norms for Neuroticism and Extraversion for school boys between the ages of 11 years and 14 years 11 months, in rural and urban secondary schools in Devon, and also to contribute data to the validation of the J.M.P.I. This enabled comparison to be made between average scores for different types of schools, and of age groups and streams within the schools. It was also possible in the case of three schools to compare individual scores on Neuroticism and Extraversion with I.Qs.

---

* It is understood that another modification of the Pintner Inventory is being employed by Mr. D. G. Mulligan, of the National Foundation for Educational Research, but this has not been published.

II.—OUTLINE OF PREVIOUS RESEARCH.

Although there is no previous research on a large scale concerning results obtained for the Junior Maudsley Personality Inventory, there are a number of reports relating to Neuroticism and Intelligence and academic achievement for adults.

The results of recent investigations suggest that there is an association between Neuroticism and Intelligence, but there is no general agreement as to whether it is a negative or positive association. Lynn (1959) found that university students tend to score higher on Neuroticism than the general population. Furneaux (1956) found that there was a positive correlation between Neuroticism and attainment in university students. Lynn and Gordon (1961) found that, though there was no significant correlation between Neuroticism and General Intelligence, subjects scoring high on Neuroticism worked faster and completed the Raven's Matrices test more quickly, and therefore, Neuroticism and vocabulary were found to be positively correlated.

Bendig (1960) found negative correlations between Neuroticism and Intelligence. Eysenck (1947) and Vernon (1937) found that neurotics were slightly below average in intelligence.

There are also some interesting findings concerning Neuroticism and age. Eysenck (1959) found that younger age groups tend to have higher scores on Neuroticism, and Gellhorn (1956) showed that there is a decline in the reactivity of the sympathetic nervous system with ageing. There is also a good deal of evidence that sympathetic reactivity underlies Neuroticism (e.g., Martin, 1961).

III.—PLAN OF THE INVESTIGATION.

(i) *Subjects and Materials.*

The J.M.P.I. was administered during the school year 1960-61 to all the 11 to 14 years 11 months boys in the secondary schools under the jurisdiction of the Local Education Authority in Exeter. This included 4 secondary modern schools and one grammar-technical school. Also a sample of the secondary schools in Devon County area was selected from the thirty-two willing to co-operate out of the total number of fifty-five. The intention was to obtain as rural a group as possible in order to compare the norms obtained for these schools with those of the urban schools. The co-operation of Exeter City and Devon County Education Authorities and of the headmasters concerned was much appreciated. This county sample consisted of three county secondary schools, one comprehensive, two grammar, and one technical high school. For the purpose of results, this technical high school was included with the grammar schools. The total number of boys covered by this investigation was 3,559.

(ii) *Experimental Procedures.*

In all the schools the J.M.P.I. was administered by teachers (although not necessarily on a form basis), except in one school where the researchers took part in the administration to two of the four groups tested. In order to standardise the administration as far as possible, instructions were issued to teachers.

These instructions emphasised that the boys should be told that the test had nothing to do with their school work, and that there were no right or wrong answers, but that it was rather an opinion poll. The boys were allowed as long as they needed to complete the test. They were also told that the teacher administering the test would not read the completed form, but merely look it over to see that each question had been answered. This resulted in extremely few tests being rejected for incompleteness.

The authors of the test had given a warning that the administration of the test by the teachers of the children might affect the validity of the responses given. Considering it was given as a group test, it is thought that bias on this account would, in fact, be reduced to the minimum. Telling the boys that the teacher would not read their answers was also directed to eliminating as far as possible bias resulting from a boy's relationship with the teacher.

It was found that the test had to be read out to those children with a reading age of less than 9 years, and teachers were asked to record the reading age of such boys on the top of the test form. A very small number were so recorded It was recognised that reading out the test items might lead to variation in interpretation of questions, but it was considered the only practicable alternative for some portions of the school population.

Absentees were not followed up as they were assumed to be randomly distributed.

Intelligence Quotients for the boys in three of the urban schools (two secondary modern and the grammar-technical schools) were obtained from the Exeter Education Department. The I.Qs. had been derived from the boys' performance in the 11-plus selection test, and they were used in this project to calculate correlation coefficients for Neuroticism and Intelligence and Extraversion and Intelligence. Unfortunately, it was not possible to persuade schools to administer a special I.Q. test for the purpose of this investigation, and therefore, I.Q. scores had to be used of all boys older than 11 years which were in an increasing degree out of date. Results for the older groups will consequently be the less reliable the more I.Q. scores vary with age.

IV.—STATEMENT OF RESULTS.

TABLE 1

ANALYSIS OF RESULTS—ACCORDING TO URBAN AND RURAL TYPES OF SCHOOL.

Type of School	Rural				Urban			
	Extraversion		Neuroticism		Extraversion		Neuroticism	
	Mean	S.D.	Mean	S.D	Mean	S.D.	Mean	S D
Grammar ……	11·9	3·34	5·99	3·48	12·11	3·26	6·42	3·47
Secondary Mod.	11·85	3·13	7·03	2·98	12·07	3·11	7·04	3·27
Comprehensive	11·62	2·99	6·68	3·27				
Maudsley Norms	12·39	3 46	7·35	3·54	12·39	3·46	7·35	3·54

(*a*) *Comparison of Urban and Rural Schools of the same type.*

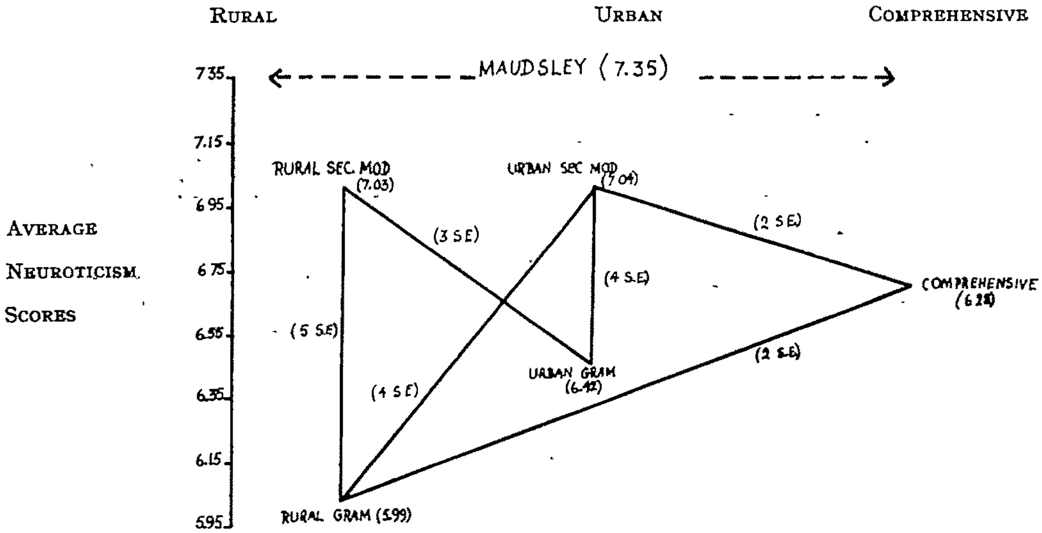
No significant differences were found between Neuroticism averages of rural and urban schools of the same type. The variance of the urban scores was slightly greater than that of the rural in the secondary modern schools. Though the urban grammar is not significantly different for Neuroticism from the rural grammar, it did have a higher Neuroticism average and a similar variance.

There were no significant differences between Extraversion averages of rural and urban schools of the same type, though there was a very slight tendency for the urban schools to have higher extraversion averages.

(*b*) *Comparison of Urban and Rural Schools of Different types.*
Neuroticism Average Scores.

DIAGRAM 1

COMPARISON OF AVERAGE NEUEROTICISM SCORES IN VARIOUS TYPES OF SCHOOLS.



NOTES :
Lines connect averages where differences are significant.
Figures on lines show number of standard error differences.

Grammar schools, both urban and rural, were lower in Neuroticism scores than secondary modern schools of any kind and than the comprehensive school. The Neuroticism average for the comprehensive school fell between the secondary modern and grammar school scores both of urban and rural schools (see diagram).

No significant differences were found between the Extraversion averages of different types of schools.

(*c*) *Comparison of Averages of "A" (more intelligent) groups and Non "A" (less intelligent) within the schools.**

In each school the combined average for Neuroticism for all the "A" boys was compared with the combined average for all the Non "A" boys.

In all the urban secondary modern schools the averages for the "A" groups for Neuroticism were significantly lower than the Non "A" averages (ranging

* "A" stream boys formed 13 per cent. of grammar school streamed classes, and 31 per cent. of modern school classes.

from 2-10×S.E.). In one of the three rural secondary modern schools the Neuroticism average for the "A" group of boys was significantly lower than the Non "A" average (5×S.E.). In the only urban grammar-technical school the "A" group Neuroticism average was very significantly lower than the Non "A" group average (4×S.E.). The three rural grammar and the comprehensive schools showed no significant differences between "A" and Non "A" group Neuroticism averages. A combined average score for Neuroticism for all the secondary modern "A" groups was found to be very significantly lower than a combined average for all the grammar school No "A" groups (more than 7×S.E.).

Being in the "A" stream of both grammar and modern schools is, therefore, usually associated with a lower Neuroticism average, though this difference is not evident among the streams in the comprehensive nor in the rural grammar schools.

There was nothing to indicate that there were any significant differences between Extraversion averages for "A" and Non "A" groups within the schools.

(*d*) *Comparison of Averages for Different Age Groups within the Schools.*

(Each school was divided into four age groups, 11-11 years 11 months, 12-12 years 11 months, 13-13 years 11 months, and 14-14 years 11 months.)

There was a tendency in nearly all the schools for the two lower age groups to have higher Neuroticism averages than the older age groups. The significant differences were as follows : In one of the three rural secondary modern schools the 11–11 years 11 months and 12–12 years 11 months averages for Neuroticism were significantly higher than the other two age groups (2 and 3×S.E. and 3 and 3×2.E. respectively) In another rural secondary modern school the 12–12 years 11 months age group Neuroticism average was signicantly higher than the 11–11 years 11 months and 13–13 years 11 months age group averages (2×S.E. for both), and the 11–11 years 11 months average for Neuroticism was significantly higher than the 13–13 years 11 months age group average In the third rural secondary modern school the 12–12 years 11 months average for Neuroticism was significantly higher than the 14–14 years 11 months average only (2×S.E.). In one of the four urban secondary modern schools the 12–12 years 11 months average for Neuroticism was significantly higher than the 11–11 years 11 months and 14–14 years 11 months averages (2 and 3×S.E.). In another urban modern school the 11–11 years 11 months and the 12–12 years 11 months Neuroticism averages were significantly higher than the 13–13 years 11 months average (both 2×S E.) In one of the three rural grammar schools the 12–12 years 11 months Neuroticism average was significantly higher than the 13–13 years 11 months Neuroticism average (2×S.E.). In another rural grammar school the 11–11 years 11 months average was significantly higher than the 14–14 years 11 months average for Neuroticism, and the 12–12 years 11 months age group average was even more significantly higher than the 11–11 years 11 months average (2×S.E. and 3×S.E., respectively). In the only urban grammar-technical school the 11–11 years 11 months average for Neuroticism was significantly higher than the 13–13 years 11 months average (2×S.E.). Only the results that were significant at the 5 per cent. level or more have been mentioned, but the same tendency for the lower age groups to have higher Neuroticism averages was also seen in the results that showed no significant differences.

TABLE 2

AGE GROUP AVERAGES FOR NEUROTICISM AND EXTRAVERSION IN DIFFERENT TYPES OF SCHOOLS.

Type of School		No. of Boys	NEUROTICISM								Extraversion							
			11 yrs. to 11 yrs. 11 ms.		12 yrs. to 12 yrs. 11 ms.		13 yrs to 13 yrs. 11 ms.		14 yrs. to 14 yrs 11 ms.		11 yrs. to 11 yrs. 11 ms.		12 yrs. to 12 yrs. 11 ms.		13 yrs. to 13 yrs. 11 ms		14 yrs. to 14 yrs. 11 ms.	
			Mean	S.D.	Mean	S.D.	Mean	S D.	Mean	S.D.	Mean	S D	Mean	S.D.	Mean	S.D.	Mean	S D.
Secondary Modern.																		
Urban	1	456	7·33	3·26	**8·00**	3·34	7·77	3·45	6·48	3·48	11·67	3·06	11·28	3·04	11·45	2·87	**12·45**	3·01
	2	476	6·77	3·30	7·21	3·34	7·08	3·12	6·84	3·00	11·95	2·73	12·15	3·10	12·10	2·94	11·79	3·27
	3	385	7·17	3·52	6·97	3·03	6·88	2·97	5·20	2·84	11·50	3·39	12·14	3·38	**13·60**	3·08	**12·59**	3·34
	4	287	7·51	2·94	**7·67**	3·31	6·63	3·08	7·03	3·01	12·31	2·74	12·12	2·77	11·44	3·14	11·97	3·47
Rural	1	159	**8·50**	2·52	**8·20**	3·39	6·68	3·12	7·76	2·73	11·20	2·52	11·58	3·35	12·09	2·77	12·16	3·43
	2	208	6·62	2·51	**7·42**	2·66	6·43	3·27	6·21	2·99	11·59	3·22	12·27	3·08	12·29	3·29	11·64	3·64
	3	196	**8·19**	2·27	**8·21**	2·40	6·81	2·73	6·55	2·92	11·62	2·24	11·88	2·25	**12·31**	3·15	11·11	2·58
Grammar :																		
Urban Grammar Tech.		578	**7·11**	3·76	**6·68**	3·34	5·84	3·49	6·42	3·35	11·96	2·94	11·99	3·60	12·38	3·35	12·03	2·97
Rural	1	70	4·00	2·20	**7·00**	2·30	5·79	3·72	6·36	3·75	12·75	2·86	12·12	2·54	12·95	3·10	11·14	4·10
	2	128	6·04	4·07	6·32	3·03	5·02	3·25	5·79	3·23	10·88	3·29	11·47	3·24	12·17	4·25	12·29	2·94
	3	192	6·13	2·74	**7·00**	3·66	5·38	4·03	6·59	3·37	11·68	3·11	12·15	2·56	11·62	3·60	12·48	3·74
Comprehensive		424	6·88	2·97	**7·05**	3·17	**6·98**	3·28	6·05	3·52	11·91	3·02	11·25	2·92	11·51	2·96	**11·94**	3·15

Scores in heavy type are significantly higher than one or more of the other age groups (see text).

As far as Extraversion was concerned, the tendency was reversed. That is, the higher age groups tended to have higher Extraversion averages than the two lower age groups. However, the only significant results showing this tendency were found in two of the four urban secondary modern schools. In one of these the 13–13 years 11 months average for Extraversion was significantly higher than all the other age group Extraversion averages (nearly 2, 2 and 2 × S.E., respectively), and in the other the 14–14 years 11 months-average was significantly higher than all the other age group Extraversion averages (2–5 × S.E. in each case).

(*e*) *Comparison of Results with Maudsley Norms.*

The Neuroticism averages for the urban and rural grammar schools were the only results that differed significantly from the results obtained by the originators of the test. The rural grammar and the urban grammar-technical averages for Neuroticism were significantly lower than the Maudsley average (4 and 3 × S.E., respectively).

(*i*) *Correlation of Neuroticism and Intelligence.*

It will be seen that in the two secondary modern schoools, Neuroticism and Intelligence are negatively correlated in every case. When individual form correlations are considered, three of the five Non "A" first forms showed significant negative correlations (two at 5 per cent. and one at 1 per cent. level); none of the second forms yielded significant correlations; two of the six Non "A" third forms showed significance (at the 5 per cent. level); and one of the four Non "A" fourth forms showed a significant negative correlation (at the 5 per cent. level).

In the grammar technical school; on the other hand, the correlations are positive in all cases but one for both the "A" and Non "A" boys. None of these correlations, however, is significant. The small size of the "A" groups should be noted.

In order to get a result for the correlation of Neuroticism and Intelligence for a random population of boys of 11–14 years over the whole intelligence range, the grammar school boys' scores were combined with the weighted scores of the two modern schools. (The weights take account of the fact that the pupils of only two of the four modern schools were tested.) The co-efficient found was $-0{\cdot}19$, which was significant.

(ii) *Extraversion and Intelligence.*

Only two significant correlations were found between Extraversion and Intelligence. The second form "A" group of the two secondary modern schools and one of the Non "A" third forms showed significant negative correlations between Extraversion and Intelligence (both at 5 per cent. level). Although over half of all the forms showed a slight tendency for Extraversion and Intelligence to be positively correlated, the correlation for the two schools combined showed a very slight negative correlation. No significant correlation was found in the grammar school between Extraversion and Intelligence. However, there was a slight tendency for Extraversion and Intelligence to be positively correlated in the 11-year and 12-year old Non "A" groups.

(*f*) *Correlation of Individual Intelligence Quotients with Individual Neuroticism Scores, and with Individual Extraversion Scores.*

TABLE 3

CORRELATION OF CO-EFFICIENTS OF NEUROTICISM-INTELLIGENCE, EXTRAVERSION-INTELLIGENCE AND NEUROTICISM-EXTRAVERSION IN TWO SECONDARY MODERN SCHOOLS COMBINED, AND IN A GRAMMAR-TECHNICAL SCHOOL.

Type of School	Year	No. in Sample (n)	Neuroticism–Intelligence		Extraversion–Intelligence		Neuroticism–Extraversion	
			Combined "A" Group	Combined Non "A" Group	Combined "A" Group	Combined Non "A" Group	Combined "A" Group	Combined Non "A" Group
Two Secondary Moderns Combined	1st	"A" 66 Non "A" 147	—0·02 —	— **—0·25**	0·03 —	— —0·03	0·07 —	— 0·02
	2nd	"A" 66 Non "A" 144	—0·13 —	— —0·15	**—0·26** —	— 0·11	—0·20 —	— —0·31
	3rd	"A" 71 Non "A" 192	—0·07 —	— **—0·21**	0·09 —	— —0·06	—0·16 —	— —0·10
	4th	"A" 66 Non "A" 114	—0·12 —	— **—0·24**	0·15 —	— —0·03	0·03 —	— 0·00
	All Years	"A" 269 Non "A" 597	—0·08 —	— **—0·18**	—0·01 —	— 0·00	—0·07 —	— —0·10
	Total	866 (b)	—0·22		—0·02		—0·08	
A Grammar-Technical	1st	97	0·08(a)		0·09		—0·09	
	2nd	"A" 14 Non "A" 106	0·09 —	— 0·07	—0·13 —	— 0·09	0·03 —	— —0·15
	3rd	"A" 18 Non "A" 123	0·06 —	— —0·03	0·13 —	— 0·01	0·30 —	— —0·25
	4th	"A" 19 Non "A" 103	0·08 —	— 0·08	—0·23 —	— 0·05	**0·57** —	— —0·03
	All Years	480 (b)	0·02		—0·02		—0·13	
Combined Modern and Grammar.	All Years	1,346 (b)	—0·19		0·01		—0·10	

Significant Correlation Coefficients in heavy type.

Notes: (*a*) First form in the grammar school was not streamed.
(*b*) "A" and Non "A" combined.

V.—DISCUSSION AND CONCLUSION.

On looking at the results superficially it seems that there is a direct relationship between scoring on Neuroticism and the level of Intelligence. For example, the secondary modern schools have significantly higher Neuroticism averages than the grammar schools and the Non "A" groups in the secondary modern and in the grammar schools tend to have higher Neuroticism averages. However, closer examination of the results suggests that it is achievement of status in the educational hierarchy that is related to scoring on Neuroticism—lower status being associated with a higher Neuroticism score. Achievement of status in the educational hierarchy depends initially upon success in the 11-plus selection test, success in most people's eyes meaning selection for a grammar school education. Secondly, acquisition of status depends upon adjustment to the academic environment of the school. Since the school setting demands academic achievement, acquisition of status in the school setting is related to academic achievement. The secondary modern school boys may score higher on average on Neuroticism than the grammar school boys in the school setting, because, in a sense, being academically inferior, they are maladjusted to the academic environment of school. The same argument applies to the Non "A" groups within the secondary modern schools and grammar schools. It is, nevertheless, interesting that, despite the fact that the grammar school Neuroticism averages are lower than the secondary modern, the secondary modern combined average of Neuroticism for all the "A" group boys is significantly lower than the grammar school combined average for Non "A's." The reason for this could be that, although the Non " A's " in the grammar school are more intelligent than the " A's " in the secondary moderns, they have relatively less status in their schools because they are not academically superior in their schools. The fact that no significant difference was found between the Neuroticism averages of "A" and Non "A" groups in the smaller rural grammar schools, whereas the difference was significant in the larger urban grammar-technical school (578 boys) could be explained by the lesser importance of status in the smaller, more personal school communities. It was interesting to find, too, that in the comprehensive school there was no significant difference between the combined "A" and the combined Non "A" Neuroticism averages or between individual "A" form and Non "A" form averages for Neuroticism ; whereas, in the secondary modern to a large extent, and in the grammar schools to a smaller extent, the "A" groups had smaller Neuroticism averages. This could be explained by the fact that although the boys may have been graded in the school itself, the initial effects of selection for a certain type of secondary education have not been at work to the same extent, and so the boys and their parents probably do not feel that the "A" and Non "A" groupings are so strongly selective.

The results of the correlations of individual Neuroticism scores and I.Qs. in the two secondary modern schools and the grammar-technical school suggest that only in the relatively lower status group (the Non "A" group of the secondary modern schools) are Neuroticism and Intelligence significantly negatively correlated. The significant negative correlation is only shown up in the larger groups, which suggests that it is not very great and is easily masked by the margin of error in the small groups. Though the correlation of Neuroticism and Intelligence in the grammar-school is not significant, it is positive. This suggests that, somewhere about this level of the educational hierarchy, the more intelligent boys score slightly higher in Neuroticism than their less intelligent school-mates. If there is greater emphasis on competitive academic position in such schools, this tendency might be explained. However, the results obtained here are inadequate as more than pointers in this direction.*

The tendency for the two lower age groups to have higher Neuroticism averages may also be due to differentials in status between the older and younger boys. As the boys move up the school, as well as gaining status for academic achievement, there will be opportunities to take responsibilities in various fields, such as becoming a school prefect. Even if there is no specific task given to each boy that increases his status, the very fact that he is a senior boy in relation to the 11 and 12-year-olds will give him increased status. The finding that the Neuroticism averages of the lower age groups tend to be higher than those of older age groups is consistent with several other findings which have suggested an overall decline in Neuroticism with ageing (Eysenck (1959), Gellhorn (1956), and Martin (1961). It would be interesting to follow this up further and investigate the level of Neuroticism among the 17 and 18-year-olds in grammar schools.

The results obtained for Extraversion do not suggest any significant relationships between it and any other factor investigated in this project, yet the tendency for the higher age groups to be more extraverted than the lower ones is interesting in that it suggests that Extraversion scores may vary slightly with age in individuals. The same may also be true of Neuroticism scores.

All the results outlined and discussed here were obtained within the school setting. Further, research using the J.M.P.I. and testing boys from similar types of schools away from the school setting might be fruitful, and help to validate the hypotheses advanced on the evidence of this project. For the findings of this enquiry lead to the conclusion that Neuroticism scores of schoolboys do reflect failure or success in the 11-plus examination, leading to a generally lower Neuroticism among grammar school pupils as a whole. Within the schools, too, selection for the "A" stream raises a boy's morale and lowers his Neuroticism score. At the same time, somewhere about the level of intelligence obtaining access to grammar schools, correlation of Neuroticism and Intelligence changes from negative—for the group of boys of lower intelligence, to positive—within the group of more intelligent pupils. Again the element of competitive spirit might be thought to explain the data. If these postulates were correct. they would imply that psychological stress in the school situation is closely associated with competitive pressure in our educational system.

---

* In this connection it is necessary to note the percentage of grammar school places in schools covered in this enquiry Exeter City 25 per cent of secondary school population ; Devon County 23 per cent

VI.—REFERENCES.

BENDIG, A. W (1960). Extraversion, neuroticism and student achievement in introductory psychology. *J. Educ. Res.*, 53, 263-267.

EYSENCK, H. J. (1947). *Dimensions of Personality.* Kegan Paul, London.

EYSENCK, H J. (1953). *The Structure of Human Personality.* Methuen, London

EYSENCK, H. J. (1959). *Manual of the Maudsley Personality Inventory.* University of London Press

FURNEAUX, W. D. (1956) *Report to the Imperial College of Science and Technology*

FURNEAUX, W. D., and GIBSON, H B (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-208.

GELHORN, E. (1956) *Physiological Foundations of Neurology and Psychiatry* University of Minneapolis Press.

LYNN, R. (1959). Two personality characteristics related to academic achievement. *Brit J. Educ Psychol.*, 29, 213-216

LYNN, R , and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit J. Educ. Psychol.*, 31, 114-203.

MARTIN, I. (1960). Chapter in *Handbook of Abnormal Psychology*, Eysenck, H. J (Ed.)

PINTNER, R., *et al.* (1938). *Aspects of Personality* World Book Co., New York and Chicago.

II.—PERSONALITY FACTORS AND ACADEMIC PERFORMANCE

By R. D. SAVAGE
(*University of New England, Australia*)*

SUMMARY Experimental studies have suggested that personality factors, in particular neuroticism and extraversion, are important determinants of academic performance. The Maudsley Personality Inventory was given to first-year university students over three years and scores on this were related to academic performance at the end of their first year. The results showed that this Australian university population had higher mean neuroticism and extraversion scores than the norms for the test. Analysis of variance and correlation techniques showed that high scores on both factors were negatively related to academic performance.

I.—INTRODUCTION.

THERE has been much speculation concerning the relationship between personality and academic success. Two personality factors frequently postulated as contributing to academic achievement are neuroticism and extraversion. These may be measured by the Maudsley Personality Inventory. Eysenckian theory predicts that good educational attainers should (1) score high on neuroticism and (2) score low on extraversion. In addition, it may be that, though successful students may have high neuroticism scores compared with the general population norms, excessively high neuroticism may be detrimental to academic success.

Severn recent studies, particularly those by Furneaux (1957) at London University, Broadbent (1958) at Cambridge, and Lynn (1959) at Exeter have confirmed the Eysenckian predictions that high drive level in terms of low extraversion are beneficial to academic success.

The present investigation was conducted on first year Arts students at New England University, Australia, between 1959-1961. It was hypothesized that (1) the higher the success of the student the lower the extraversion score; (2) the mean neuroticism score of the group will be higher than that for a general population; (3) the higher the academic success of the students, the lower the neuroticism score. The third prediction stems from a development of Eysenck's basic postulate to allow for a U-shaped relationship between neuroticism and performance. An optimum level of performance may be related to an optimum drive level for a task of given difficulty.

II.—METHOD.

One-hundred and sixty-eight male and female students entering the Arts Faculty of New England University in 1959, 1960, 1961, were given the Maudsley Personality Inventory (Eysenck, 1957) in their second term. The neuroticism and extraversion scores of each student were related to the results obtained in the annual examinations. All students in this study attempted four first-year subjects within the Arts Faculty curriculum. The students were divided into five groups according to the number of passes out of four obtained in the final examinations: group 1—no passes, group 2—one pass and so on. Finally, analyses of variance were performed on the neuroticism and extraversion scores of the five groups. In addition, mean neuroticism and extraversion scores were correlated with academic success.

III.—RESULTS.

The mean neuroticism score for this university population was 27·7. The neuroticism scores of groups 1, 2, 3, 4 and 5 were subjected to analysis of variance. The means for the groups were 40, 31·2, 27·1, 28·1 and 26·2 respectively. The variances between groups were significant at the 1 per cent. level (Table 1). There was a negative correlation of ·9 which is significant at the 5 per cent. level, between the mean neuroticism score for each group and academic success.

The mean extraversion score for the population was 27·3. Analysis of variance of the extraversion scores showed mean scores of 36·17, 35·14, 28·18, 25·5 and 26·1 for groups 1, 2, 3, 4 and 5 respectively. The variances between groups were significant at the 1 per cent. level (Table 2). The correlation between the mean extraversion score for each group and academic success (– ·9) was significant at the 5 per cent. level.

TABLE 1

ANALYSIS OF VARIANCE OF NEUROTICISM SCORES.

	S. Sq.	d.f.	M S V.	F.
Total Variance	16843	167	—	—
Variance between groups ..	1416	4	354	3·74
Variance within groups	15427	163	94·6	—

TABLE 2

ANALYSIS OF VARIANCE OF EXTRAVERSION SCORES.

	S. Sq.	d.f.	M.S.V.	F
Total Variance	13518	167	—	—
Variance between groups ..	1608	4	402	4·9
Variance within groups	12110	163	84·3	—

IV.—DISCUSSION.

The results of this investigation show that neuroticism and extraversion scores on the Maudsley Personality Inventory are significantly related to academic performance. The mean neuroticism and extraversion scores for this university population were significantly higher than the English or American norms for the M.P.I. (Eysenck 1957) and comparable to those found by Lynn (1959) for English university students.

The positive association between neuroticism and academic failure does not follow the results of Furneaux (1951), (Lynn 1959) and others. Groups 3 and 4, who passed three and four subjects out of four had a mean neuroticism score significantly above the norm for the test. Groups 1, 2 and 3, however, had significantly higher neuroticism scores than the other two groups. These higher scores tended to impair performance. It may be suggested than an optimum level of neuroticism exists for successful academic achievement and that there is a U-shaped relationship between the variables.

The relationship found between extraversion and academic performance confirmed the results of Furneaux (1957), Broadbent (1958) and Lynn (1959), with British university subjects. The most successful group was that with the lowest extraversion scores and the other groups ordered in the predicted direction.

The interesting problem still exists as to the interaction of neuroticism and extraversion in relation to academic performance. Further investigation of this is necessary. The present results suggest that academic performance X may be related to a neuroticism score A plus an extraversion score B, or in neuroticism score B plus an extraversion score A.

V.—REFERENCES.

BROADBENT, D. E. (1958). *Perception and Communication.* London: Pergamon Press.

EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria* London: Routledge and Kegan Paul.

FURNEAUX, W. D. (1957). *Report to Imperial College of Science and Technology*

LYNN, R. (1959) Two personality characteristics related to academic performance *Brit. J. Educ. Psychol.*, 31, 3, 213-216.

III.—CULTURAL AND SEX DIFFERENCES IN EXTRAVERSION AND NEUROTICISM REFLECTED IN RESPONSES TO A CHILDREN'S PERSONALITY INVENTORY*

By C. G. COSTELLO and H. M. BRACHMAN
(*Munroe Wing, Regina General Hospital*)

SUMMARY. A new children's personality inventory—the J M P.I.—designed to measure extraversion and neuroticism was administered to 261 male and 248 female Canadian high school children and to 304 male and 230 female English comprehensive school children

The only difference in mean scores obtained in comparison of cultures and sex was in mean Neuroticism score for sex. Both the Canadian and English girls obtained higher mean N scores than Canadian and English boys.

Item analyses suggested differences between males and females and between Canadian and English children in the expression of extraversion and neuroticism.

There is a suggestion that acquiescence, as measured by the number of 'Same' responses, is greater among English school children than among Canadian school children.

I.—INTRODUCTION.

In their discussion of some J.M.P.I. data, Furneaux and Gibson (1961) state that "there appears to be some reason for believing that the results of an item analysis of inventory material of this kind depends in part on the sex of the group within which the analysis is conducted." Carrigan (1960) has pointed out discrepancies between factorial studies based on male and female subjects which suggest a qualitative sex difference in the extraversion-introversion dimension of personality. There is also the problem of cultural differences in relation to personality inventories. Bronzaft, Hayes, Welch and Koltuv (1960) have questioned the suitability for Americans of the Maudsley Personality Inventory (1959), which has English norms

This study investigates possible sex differences and differences between Canadian and English children in extraversion and neuroticism as measured by the J.M.P.I.

II.—METHOD.

The J.M P.I was administered under group conditions to 261 male and 248 female Canadian high school children between the ages of 14 and 16, and to 304 male and 230 female English comprehensive school children between the ages of 14 and 16 The procedure followed in the administration of the inventory was identical to that of Furneaux and Gibson (1961).

III.—RESULTS.

The distribution data for the Canadian sample, our English sample and the Furneaux and Gibson English sample are presented in Table 1.

---

* The authors wish to express their indebtedness to the teachers at Sedgehill School, London, England, and Martin Collegiate and Sheldon-Williams Schools, Regina, Canada, for their assistance in the collection of the data. They also wish to express their gratitude to Mr. Ron Boychuk for his computational work.

TABLE 1

J.M.P.I. DISTRIBUTION DATA FOR TWO ENGLISH SAMPLES (N=156, N=534) and One CANADIAN SAMPLE (N=509).

		Extraversion	Neuroticism
Furneaux and Gibson (1961)	Mean SD N	12·39 3·46 156	7·35 3·54 156
Costello and Brachman (Canadian Sample)	Mean SD N	12·05 3·61 509	5·90 2·84 509
Costello and Brachman (English Sample)	Mean SD N	12·16 3·35 534	6·30 3·17 534

The differences between the three mean Extraversion scores are not statistically significant. The mean Neuroticism score for the Canadian sample is significantly lower than the mean Neuroticism score of the Furneaux and Gibson sample (t=4·68, df=663, $p < \cdot 01$). The mean Neuroticism score for our English sample is significantly lower than the mean Neuroticism score for the Furneaux and Gibson English sample (t=3·33, df=1041, $p < \cdot 01$). The age range of the Furneaux and Gibson English sample was 9–14 years and it may be this difference in age range between their sample and our two samples that accounts for the difference in mean Neuroticism score. If this is so, it would indicate the need for norms on personality inventories to be established for different age groups. Data in this connection are being collected at present.

The correlations between the Extraversion and Neuroticism scales for the three samples are as follows: Furneaux and Gibson English sample r= – ·10; Canadian sample r= – ·179; English sample r= – ·206. The two negative correlations in the present study are statistically significant($p < \cdot 01$) and are somewhat higher than that obtained by Furneaux and Gibson. The significance of the correlations suggests that the scales are in need of purification.

The twenty-two items directed to the measurement of extraversion had, in Furneaux and Gibson's English sample, correlations ranging from 0·29 to 0·80 with the scale they defined, mean 0·52. In our English sample, these items had correlations with the E scale ranging from 0·22 to 0·64, mean 0·42. In the Canadian sample, these items had correlations with the E scale ranging from 0·24 to 0·58, mean 0·42.

The twenty-two items directed to the measurement of neuroticism had, in Furneaux and Gibson's sample, correlations with the N scale ranging from 0·28 to 0·64, mean 0·45. In our English sample these items had correlations with the N scale ranging from 0·22 to 0·62, mean 0·47. The correlations of these items with the N scale in our Canadian sample ranged from 0·31 to 0·68, mean 0·48.

The range of correlations between items and the scale they are designed to measure is satisfactory in all three samples. Some of the items also have sizeable correlations with the scale they are not designed to measure, however, and this may necessitate their omission.

The significance of the differences between the correlations obtained by the item analyses were investigated to reveal possible sex and cultural differences with the following results.

Four of the Extraversion items had significantly different correlations with the E scale when the Canadian sample was compared with the English sample.

Comparing the same two samples nine of the Neuroticism items had significantly different correlations with the N scale.

Within the Canadian sample, five of the Extraversion items and four of the Neuroticism items revealed significant sex differences. Within the English sample seven of the Extraversion items and six of the Neuroticism items revealed significant sex differences.

Acquiescence

The response to each item on the Junior M.P.I. is either ' Same ' or ' Different,' indicating agreement or disagreement with the statement made in the item. A measure of acquiescence was obtained by counting for each subject the number of ' Same ' responses he or she gave. The results for our two samples are given in Table 2.

TABLE 2

NUMBER OF " SAME " RESPONSES ON THE J.M.P.I. FOR CANADIAN CHILDREN AND ENGLISH CHILDREN.

Canadian	Male	Mean SD	17·90 3·49
	Female	Mean SD	19·08 3·13
English	Male	Mean SD	20·00 4·15
	Female	Mean SD	21·26 4·04

The girls in both the English and Canadian sample had significantly higher acquiescence scores than the boys. This may, however, be an artifact. In Table 3 the distribution data for the E and N. scales has been broken down to reveal sex differences.

There is no significant difference between the sexes in mean E score but the females in both the English and Canadian samples have a higher mean N score than the males. The higher acquiescence score of the girls may be related to their higher Neuroticism scores since the keyed responses that make up the N scale are 15 ' Same ' and 7 ' Different.'

The English boys and girls had significantly higher acquiescence scores than the Canadian boys and girls. Since there was not a significant difference between the English and Canadian children in mean N score, the difference in mean acquiescence score between the two cultures cannot be accounted for in terms of differences in Neuroticism.

TABLE 3

J M.P.I. DISTRIBUTION DATA FOR CANADIAN BOYS AND GIRLS AND ENGLISH BOYS AND GIRLS.

			Extraversion	Neuroticism
Canadian	Male	Mean	12·18	5·28
		SD	3·149	2·598
	Female	Mean	12·13	6·56
		SD	3·257	2·561
English	Male	Mean	12·26	5·99
		SD	3·425	3·074
	Female	Mean	12·04	6·72
		SD	3·233	3·252

IV.—REFERENCES.

BRONZAFT, A., HAYES, R., WELCH, L , and KOLTUV., M. (1960). Relationships between extraversion, neuroticism and ascendance *J. Psychol* , 60, 279-285.

CARRIGAN, PATRICIA M. (1960) Extraversion-introversion as a dimension of personality : a reappraisal. *Psychol. Bull.*, 57, 329-360

EYSENCK, H J (1959). *Manual of the Maudsley Personality Inventory.* London : Univ. of London Press.

FURNEAUX, W. D. and GIBSON, H. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol* , 31, 204-207.

(*Manuscript received*, 26*th January*, 1962.)

CHILDREN'S EVALUATIONS OF THE CHARACTERISTICS OF THE GOOD TEACHER

By PHILIP H. TAYLOR
(University of Leicester Institute of Education)

SUMMARY. Over 800 children from primary and secondary schools ranked items descriptive of a ' good ' teacher in four scales. The items of each scale were representative of statements made by children in short essays about 'A Good Teacher.' The four scales were so organised as to sample three hypothesized areas of the good teacher's class-room behaviour: discipline, teaching and personal qualities; and to bring together in the fourth scale these three areas for comparative evaluation.

A sample of teachers and training college students also completed the comparative scale.

Together with the scales a check-list of twenty pairs of items was administered to the children. Children from junior schools also completed a scale concerned with the ' good ' teacher's class-room organisation.

Differences were found between children at different stages of education, between children and teachers, between teachers and students, and within teachers, between graduate and non-graduate teachers. The reasons for these differences are discussed.

I.—INTRODUCTION.

THIS enquiry originated as an aspect of the work of twenty-one teachers studying part-time for a Diploma in Education. It arose out of a discussion of the research on teacher characteristics reported by Witty (1947) and commented on by Vernon (1953).

It was clear from a survey of the literature that few evaluations of teacher characteristics made by children have been undertaken in this country. It was also clear that recommendations to do this were by no means new. James Ward (1926), lecturing towards the end of the nineteenth century on the application of mental science to education, had suggested to teachers, " Surely one of the first steps towards the understanding of the young is to know how they regard us." Evans (1954) has pointed out that studies of the kind suggested were undertaken in this country and elsewhere in the last decade of the nineteenth century.

However, by the mid-thirties of this century, such studies had petered out. Children's evaluations of teachers do crop up, but as incidental rather than central themes in researches. For example, Morris (1955, 1958) studying the moral values of adolescents incidentally studied the changes in attitudes to teachers using a role comparison method and found, " a tendency for older pupils to choose more impersonal roles as like the good teacher." Similarly, Allen (1959) studying attitudes to school of secondary modern children found that both boys and girls valued highly teachers being competent, teachers really teaching, lessons being interesting, teachers joking, taking a joke, being friendly and approachable, and possessing understanding. He found that boys valued more highly than did girls a teacher's ability to teach, and that girls valued more highly the teacher's general sociability. He also found that both boys and girls valued most highly the general competence of the teacher.

The only study in this country directly concerned with children's evaluations of teacher's behaviour is that of Hollis (1935) who asked 8,043 children of different ages in both mixed and single sexed schools to rank seven statements descriptive of a teacher's behaviour. His findings were that the quality of 'explaining difficulties patiently is conclusively the most popular. 'Being friendly and sympathetic' was ranked second, with 'just and fair' third. Fourth and fifth were 'humour ,' and 'allowing pupils to ask plenty of questions' The teacher 'having wide interests' was ranked sixth, while 'discipline' was ranked as of least importance.

Whenever the question of children rating teachers has been raised (Vernon 1953, Bush 1954, Coffman 1954, and Gage, *et al.*, 1955) it has been acknowledged as a reliable method of judgment. Rather more enquiries in this field have been carried out in the United States. Allen (1960) has effectively summarised and commented on them.

II.—OUTLINE OF THE ENQUIRY.

Stage 1. Some 866 children in twelve junior schools, 401 in four secondary modern schools and 112 children in one grammar school were asked to write two short essays on 'A Good Teacher' and 'A Poor Teacher.' Between twenty and thirty minutes were allowed for both essays.

The 1,379 essays were then analysed for content by twenty-one teachers. Each teacher analysed a separate batch of essays. Every independent statement made about the 'good' and the 'poor' teachers' behaviour was then allocated to one of four categories considered to be descriptive of mutually exclusive areas of a teacher's class-room behaviour. This was done separately for essays about the 'good' and the 'poor' teacher. The four categories were: Teaching (T), Discipline (D), Personal Qualities (P) and Organisation (O). The meaning to be attached to these categories was then clarified in general discussion after a sample of the esssays had been read by the teachers concerned in the enquiry.

At the same time as the essays were being written by the children and analysed by the teachers, a sample of eight-four teachers (fifty-three in junior schools, and thirty-one in secondary schools) were asked to state five characteristics which most clearly represented the behaviour of the 'good' teacher.

Stage 2 What follows, reports only the use made of the analysis of the essays 'A Good Teacher' written by the children.

The analysis of the essays into statements allocated to one of the four areas, T, D, P and O, provided items for the construction of five scales, each of six items and a check-list of twenty pairs of words and phrases. Scale A contained the two items which occurred most frequently within the three major areas, T, D and P Scales B, C, D and E contained the most frequently occurring items within the four categories T, D, P and O.

The wording of the items in each scale was as typical of children's statements as possible. The items were arranged in random order within each scale and each scale was prefaced by precisely the same instructions (Appendix A) to rank the items in order of importance from one to six

The instructions accompanying the check-list of words and phrases were to tick only those words or phrases considered by the child to be descriptive of a good teacher.

The scales and check-list in booklet form were then administered to 897 children: 500 in the fourth year of the junior school, 230 in the second year and 167 in the fourth year of the secondary school.

Fourteen junior schools, six secondary modern schools and two grammar schools co-operated. The schools came from both urban and rural areas and without exception came under a local authority which organised its schools in the conventional manner and used 11+ selection.

It should be noted that the schools which co-operated were not a random selection of all the schools within the area Both the schools and the classes within them were used because they were accessible to the teachers carrying out this enquiry.

Scale A was completed correctly by 131 out of the 135 teachers who were approached, and by forty-three first-year training college students from a mixed college.

III.—RESULTS.

(i) *Analysis of Scale A.*

Scale A consisted of the following items :

*(*a*) A good teacher is fair and just about punishment and has no favourites.
(*b*) A good teacher explains the work you have to do and helps you with it.
(*c*) A good teacher is patient, understanding, kind and sympathetic.
(*d*) A good teacher is cheerful, friendly, good-tempered and has a sense of humour.
(*e*) A good teacher is firm and keeps order in the classroom.
(*f*) A good teacher encourages you to work hard at your school work.

The method adopted for the analysis of the rankings of Scale A was to sum the inverse of the rankings for the pairs of items considered to be representative of the three areas of the teacher's behaviour : T, D and P.

TABLE 1

SCALE A. SUMS OF RANKING RELATED TO CATEGORIES OF A ' GOOD ' TEACHER'S CLASSROOM BEHAVIOUR.

Scale A		Children			Teachers						Students	
items	Area	4y. J	2y. S	4y. S	Jun.	Sec.	Men	Women	N.Grd'	Grad.	Jun.	Sec.
b+f	T	4187	1929	1385	457	390	502	345	658	189	127	130
a+e	D	3701	1624	1134	442	397	479	360	667	172	93	172
c+d	P	2612	1277	988	571	494	636	429	880	185	179	202
	N	500	230	167	70	61	77	54	105	26	19	24
Percentage Distribution	T	39·9	39·9	39·5	31·1	40·4	31·1	30·4	29·8	34·6	31·8	25·8
	D	35·2	33·6	32·3	30·1	31·1	29·6	31·8	30·3	31·5	23·3	34·1
	P	24·9	26·4	28·2	38·8	38·6	39·3	37·8	39·9	33·9	44·9	40·1
Value of X^2 and Level of P.		5·6 2·6 N S. N.S. 17·7 ***			2·0 N.S.		1·5 N S.		7·6 *		12·9 **	

76·9 per cent of the graduate teachers were teaching in grammar schools.

* These letters are for textual reference only and were not printed on the original scales

As Table 1 indicates, children at all stages gave most weight to the ' good ' teacher's teaching. They accord this area about 40 per cent. of their rankings. They gave least weight to the ' good ' teacher's personal qualities. Teachers on the other hand gave most weight to the ' good ' teacher's personal qualities, according this area some 39 per cent. of their ranking. Teachers gave about equal weight to the ' good ' teacher's teaching and discipline.

Differences between children were found only between junior children and fourth year secondary children ($X^2=17 \cdot 7$, $P > 0 \cdot 001$). Junior children placed rather greater emphasis on the ' good ' teacher's discipline, in particular on his being fair and just about punishment and having no favourites. Fourth year secondary children, on the other hand, placed rather more emphasis on the ' good ' teacher's personal qualities, particularly on his being cheerful, good-tempered and having a sense of humour.

Teachers in both the junior and secondary schools differed significantly from children at the junior and at the secondary stage, respectively, in the emphasis they placed on each of the three areas ($X^2=54 \cdot 4$ and $60 \cdot 7$, respectively, $P > 0 \cdot 001$ in each case). Teachers emphasised to a greater extent than children the ' good ' teacher's personal qualities, and to a lesser extent his teaching and discipline.

Men and women teachers did not differ significantly, nor did teachers in primary and secondary schools, nor were differences to be found when teachers were grouped into categories of less than five year's experience, between 5 and 25 years' experience, and over 25 years' experience. ($X^2=5 \cdot 31$, N.S.) Teachers in all groups gave most weight to the ' good ' teacher's personal qualities, although there was a tendency for longer serving teachers to give greater weight to the ' good ' teacher's discipline at the expense of his teaching. Graduate and non-graduate teachers differed at the 5 per cent. level. The former placed greatest emphasis on the ' good ' teacher's teaching and the latter on his personal qualities.

The small sample of training college students tended to take up a more exaggerated position than did practising teachers. The group training to become junior school teachers placed greater emphaisis on the ' good ' teacher's teaching while students training to become secondary school teachers placed greater emphasisjon the ' good ' teacher's discipline. Students, no matter what type of school they were training for, placed the greatest emphasis on the ' good ' teacher's personal qualities; and men students did not differ significantly from women students.

Differences between boys and girls at the same stage, and between boys and girls at different stages did not concern the emphasis placed by all children on the ' good ' teacher's teaching. Differences concerned the relative emphasis on the discipline and personal qualities of the ' good ' teacher. Fourth year secondary boys placed significantly greater emphasis on the ' good ' teacher's personal qualities than girls, who emphasised the 'good' teacher's discipline. Boys and girls at the primary and at the second year secondary stage did not differ. However, both primary and secondary boys differed significantly from fourth year secondary boys; the former emphasising discipline to a greater extent than personal qualities and conversely for the latter. ($X^2=16 \cdot 9$ and $19 \cdot 7$, respectively, $P > 0 \cdot 001$ in both cases).

Jniuor girls differed significantly from second year secondary girls. ($X^2=6 \cdot 3$, $P > 0 \cdot 05$.) The former emphasised the ' good ' teacher's discipline to a greater extent than the latter. No significant differences were noted between either junior or second-year secondary girls and fourth year secondary girls.

Differences between children in different types of school were found only between children at the junior stage in streamed and unstreamed schools ($X^2=18\cdot7$, $P>0\cdot01$). Children in unstreamed junior schools were more concerned with the 'good' teacher's personal qualities and those in streamed schools were more concerned with his discipline Children in both types of school were equally and most concerned with the 'good' teacher's teaching.

(ii) *Analysis of Scale B.*

The analysis of Scales B, C, D and E was based on the over-all rankings obtained from the mean sums of the original rankings. Throughout Spearman's *rho* is used.

Scale B was constructed to sample the disciplinary area of the 'good' teacher's class-room behaviour. Discipline was conceived of as being concerned with both rewards and punishments, and with the manner as well as the method of discipline. Scale B items were:

(*a*) A good teacher is firm and keeps order in the classroom.
(*b*) A good teacher is fair and just about punishment.
(*c*) A good teacher praises you for behaving well and for working hard.
(*d*) A good teacher has no favourites.
(*e*) A good teacher lets you have some of your own way.
(*f*) A good teacher uses the cane or strap when necessary.

There was agreement at the 1 per cent. level ($r_s=+\cdot9$) between all children at different stages in their education in their ranking of the items. All children ranked first either item (*a*) 'A good teacher is firm,' or item (*b*) 'A good teacher is fair and just about punishment.' They ranked sixth item (*e*) 'A good teacher lets you have some of your own way.'

Boys at all stages were in agreement at the 1 per cent. level in this ranking of the items but girls were not ($r_s=+\cdot9$ or over). There was not significant agreement between fourth year secondary girls and younger girls. They differed most in their ranking of items (*a*) and (*d*); fourth-year secondary girls were more concerned with the 'good' teacher having no favourites (ranked first) rather than being firm (ranked third), whilst the converse was true for younger girls

(iii) *Analysis of Scale C*

Scale C was constructed to sample the 'good' teacher's teaching as seen by children. The items were:

(*a*) A good teacher encourages you to work hard at school
(*b*) A good teacher explains the work you have to do and helps you with it
(*c*) A good teacher knows a great deal about the subject he is teaching
(*d*) A good teacher gives interesting lessons
(*e*) A good teacher gives you time in the lesson to finish your work
(*f*) A good teacher marks your work regularly and fairly

No significant agreement was found between secondary children and junior children Between secondary children agreement was high ($r_s=+\cdot94$, $P>0\cdot01$)

Junior children ranked first ietm (*a*), the 'good' teacher encouraging them to work hard, while secondary children ranked this item third or fourth, ranking item (*b*), the 'good' teacher explaining the work first. All children ranked item (*c*), the 'good' teacher knowing his subject second; and sixth, item (*e*), the timing of the lesson.

Agreement between boys at different stages was significant at the 5 per cent level or above but between girls there was agreement only at the secondary stage. ($r_s = +\cdot 94$, $P > 0\cdot 01$.) A ' good ' teacher knowing his subject and giving interesting lessons were ranked higher by secondary girls than by junior girls. Junior girls ranked the ' good ' teacher's encouragement to work higher than secondary girls : first as against third and fourth, respectively.

(iv) *Analysis of Scale D.*

Scale D was constructed to sample the ' good ' teacher's personal qualities ; his manner, temperament and appearance, and was as follows :

(*a*) A good teacher is cheerful and good-tempered.
(*b*) A good teacher looks nice and dresses well.
(*c*) A good teacher is well-mannered and polite.
(*d*) A good teacher is patient, understanding, kind and sympathetic.
(*e*) A good teacher has a sense of humour.
(*f*) A good teacher is friendly with children in and out of school.

Substantial agreement ($r_s = +\cdot 94$, $P > 0\cdot 01$) was found between junior and second year children secondary but not between junior and fourth year secondary children ($r_s = +0\cdot 60$, N.S.), or between second and fourth year secondary children ($r_s = +0\cdot 77$, N.S.).

All chileren agreed, in ranking item (*b*), the ' good ' teacher's appearance as least in importance and in placing item (*f*), the ' good ' teacher being friendly, first or second. The main difference was in the emphasis fourth year secondary children placed on the ' good ' teacher being cheerful and good-tempered ; item (*a*). This item was ranked second by fourth year secondary children but only fourth by younger children.

Boys at all stages agreed in their ranking at the 5 per cent. level or above, whereas girls did not show significant agreement in their ranking. Both groups of secondary girls differed from junior girls in ranking item (*a*), the ' good ' teacher being cheerful and good-tempered, much more highly (1st or 2nd). Junior girls enphasised the manners and politeness of the ' good ' teacher, placing this item, (*c*), first.

Second year secondary girls differed from fourth year girls most in their ranking of item (*c*). The second year girls rank it fifth and the fourth year, third.

(v) *The Analysis of Scale E.*

Scale E was analysed for junior school children only. The items were :

(*a*) A good teacher makes certain that the classroom is tidy and attractive.
(*b*) A good teacher has work ready for you as soon as you get into the classroom.
(*c*) A good teacher makes sure you have the pens, paper and books you need.
(*d*) A good teacher lets children help to give out books, pencils and paper.
(*e*) A good teacher knows where to find the things he wants.
(*f*) A good teacher is able to organise all kinds of activities in the classroom.

Within this area of organisation there was marked agreement between boys and girls ($r_s = +0\cdot 99$, $P > 0.01$). They both ranked first the ' good ' teacher's making sure they had materials to work with (item (*c*)) and last the help with the distribution of materials which a ' good ' teacher may organise through children.

(vi) *The Analysis of the Check-list.*

The check-list (Appendix B) was completed correctly by 880 of the 897 children. Rankings for all groups were significantly related at the 5 per cent. level or above The items carrying an asterisk were checked by 50 per cent or more of the children within each group. Only one item, ' joins in ' was checked by 50 per cent. or more of the children in all groups. The items ' old,' ' woman,' ' is like my father,' ' is like my mother,' ' fat ' and ' doesn't join in ' were checked by fewer than 15 per cent. of the children in each of the groups.

IV.—DISCUSSION AND CONCLUSIONS.

The analysis of the composite Scale A constructed to sample three areas of the ' good ' teacher's classroom behaviour indicated that children in general evaluated most highly the ' good ' teacher's teaching. This is in keeping with the findings of Hollis (1935) and Allen (1959) Children in general evaluated least highly the ' good ' teacher's personal qualities

The majority of teachers, on the other hand, evaluated most highly the ' good ' teacher's personal qualities particularly his patience, kindness, sympathy and understanding, as did training college students. The minority of graduate teachers did not They emphasised the ' good ' teacher's teaching.

Evaluations of the kind made by children and teachers in this enquiry are, no doubt, determined by the psychological location of the rater, that is by his personal needs, his perception of educational objectives and by his status—as a pupil, a teacher with particular background and training, or as a student.

The majority of teachers sampled in this enquiry would seem to be depending for their evaluation of a ' good ' teacher on the quasi-psychological view of the child current in education, especially in teacher training, which stresses the need the child has for the sympathetic, kindly understanding of a patient, cheerful, good-tempered, adult and which tends to play down the need the child may have to be taught and to learn.

The training college students would appear to make their evaluations similarly, but with greater emphasis. The minority of graduate teachers' evaluations emphasise the ' good ' teacher's teaching. Their view is no doubt related to their subject orientated education and training which emphasised the qualities they rate highly : being taught and learning.

Children in streamed junior schools where formal class teaching tends to be common emphasised the ' good ' teacher's discipline, particularly his being firm and keeping order, rather than his personal qualities It is probably that this behaviour on the part of the class teacher in the streamed school is seen by the children as essential if their need to be taught is to be met. Similarly, in the unstreamed junior schools, children emphasise the ' good ' teacher's personal qualities, particularly his patience, sympathy, kindness and understanding, which is no doubt required of the teacher in the informal situation demanded in teaching children of widely differing abilities.

All children of both sexes and in all schools evaluated most highly the ' good ' teacher's teaching which they probably perceive as a means to the satisfaction of a need they have in our society : to be taught and to learn.

However, children at different levels of maturity have different needs. Thus, fourth year secondary school children, more especially the boys, in emphasising to a greater extent than younger children, the ' good ' teacher's personal qualities, particularly his cheerfulness, good temper and sense of humour are perhaps underlining their need to be taught by teachers who are, as Morris (1958) puts it, " equitable."

The analyses of the remaining scales indicate that the item ranked sixth, of least importance in each scale, is common for all children at all stages. Thus they rank sixth the characteristic of the 'good' teacher's discipline which, suggests his being permissive; of his teaching, his timing of lessons and of his personal qualities, his appearance and dress.

General agreement on items ranked first was not found. There were differences between sexes and within sexes, between stages. However, the general picture suggests that within the area of the 'good' teacher's discipline, the ability to be firm and keep order together with his fairness about punishment, particularly his having no favourites in this matter, are ranked highly by many children. Within the area of teaching, the 'good' teachers' ability to explain and to help with school work together with his knowledge of the subject and his readiness to encourage learning, are ranked highly by many children.

Friendliness and cheerfulness and good temper are ranked most highly within the area of personal qualities by both boys and girls in the fourth year of the secondary school and by girls in the second year. Junior school children and second year secondary school boys ranked most highly the good manners and patient, kindly understanding of the 'good' teacher.

The analysis of the check-list suggests that a significant characteristic of the 'good' teacher in the eyes of children may well be his readiness to join with them in their activities.

The analysis also suggests that children's stereotype of the 'good' teacher may well be that of a young, married man with children who tends to give little homework and does not use the cane. The analysis would seem also to suggest that children may tend to reject as 'good' teachers women, elderly teachers, and those teachers inclined to behave towards them as their parents might.

ACKNOWLEDGEMENTS.—The author wishes to thank the students studying for the Diploma in Education who were in their first year in 1961 and who carried through this enquiry with enthusiasm. They are: J. C. P. Hale, K. D. Driver, F. Garbutt, D. Hall (Miss), L. W. Hopkins, K. G. Kent, N. Lowson, D. O. Lovegrove (Miss), A. D. McIntyre, M. Mobbs (Mrs.), K. Moone, E. C. Nickerson, W. H. S. Norton, C. W. Pugh, E. Quartly (Miss), E. S. Robinson, M. A. Roughton, W. R. Stobart, D. B. White, H. K. Williams, and C. R. Wylie.

V.—APPENDIX.

A. *Preface to Scales A to E*: Read the following statements carefully two or three times. When you have read them carefully put 1 in the box by the side of the statement which *you* think is most important; then put 6 in the box by the side of the statement which you think is *least* important. Put 2 in the box by the side of the statement which you think is next in importance to 1, and so on with 3, 4 and 5.

B. old.
young*
married*
single
man*
woman
fat
thin
has own children*
has no children of own
is like my father
is not like my father
joins in*
doesn't join in
is like my mother
is not like my mother
uses the cane
doesn't use the cane*
gives little homework*
gives a lot of homework

VI.—References.

Allen, E. A. (1959). *Attitudes to School and Teachers in a Secondary Modern School.* M.A. (Ed.) Thesis (unpublished), University of London.

Allen, E. A. (1960). Attitudes of children and adolescents in school. *Ed. Res.*, 111, 1, 65-80.

Bush, R. N. (1954). *The Teacher Pupil Relationship.* New York: Prentice Hall.

Coffman, E. W. (1954). Determining students' concept of effective teaching from their ratings of instructors. *J. Educ. Psychol.*, 45, 5.

Evans, K. M. (1954). The qualities of a good teacher. *Educ. for Teaching*, 35.

Gage, N. L., Leavitt, G. S., and Stone, G. C. (1955). Teachers' Understandings of their pupils and pupils' rating of their teachers. *Psychol. Monog.*, 69, 21.

Hollis, A. W. (1935). *The Personal Relationship in Teachings.* M.A. Thesis (unpublished) University of Birmingham.

Morris, J. F. (1955). *A Study of Value-judgments in Adolescents.* Ph.D. Thesis (unpublished), University of London.

Morris, J. F. (1958). Symposium: the development of moral-values in children. II.—The development of adolescent value-judgments. *Brit. J. Educ. Psychol.*, 28, 1-14.

Vernon, P. E. (1953). The psychological traits of teachers. Section 1, Chapter 3 in *Year Book of Education*, 1953. London: Evans' Bros.

Ward, J. (1926). *Psychology Applied to Education*: Lectures given at Cambridge in 1880. Cambridge University Press.

Witty, P. (1947). An analysis of the personality traits of the effective teacher *J. Educ. Res.*, 30, 662-71.

(*Manuscript received, 1st March,* 1962.)

EXAMINING ESSAYS FOR ELEVEN PLUS CLASSIFICATION

BY D. YOUNG
(*Educational Psychologist, Barnsley Education Committee*)

SUMMARY. A procedure for overcoming the marking and administrative problems of using the essay in selection examinations is described. It combines assessments of children by their own teachers with an external essay examination. The difficulties of this are overcome (1) by a system of splitting the essays by sex and title and by further sub-dividing for convenience of size of batches and (2) by using prepared tables which impose a mean of 100 and a standard deviation of 15 on these batches. Certain weaknesses of the usual marking routines are examined and evidence offered for the need for variety of titles if children are to be assessed fairly.

Validity as measured against success in grammar school examinations of varied content is as high as that for objective tests of verbal reasoning, English and arithmetic.

I.—INTRODUCTION.

IT has been argued that the use of objective tests in English leads to a neglect of writing of consecutive prose and that the addition of an essay or essays to the testing would help to combat this tendency and would lead to a higher standard of written English. The varying standards of markers and the inconsistencies of markers and children have long been recognised, and in recent years, schemes for dealing with these difficulties have emerged. These usually take the form of an examination consisting of at least three essays per child, each of which is marked three times by a team of markers. This is the minimum if a really reliable score is desired. (Vernon, 1957).

There are obvious administrative difficulties to be overcome in arranging three examination sessions, plus extra sessions for absentees, and in dealing with the set of nine marks for each child. Even if the borderzone children's essays alone are marked there is bound to be some cumbersome office work. Furthermore, it is unfortunately true that marking policy becomes common knowledge and it might be doubted that confining the marking to the borderzone would have the required effect at other levels of ability. Ideally, assuming the fact of backwash and the efficiency of the remedy, there can be no doubt that the remedy should be applied in such a way as to affect all those children who are capable, with appropriate teaching, of writing an essay.

Yates and Pidgeon (1957a and 1957b) have offered another way of combating backwash, i.e., by the introduction into objective tests of items which demand creative responses. These have shown satisfactory reliability and validity but their relationship to essay writing is uncertain. Criticism of objective tests tends to be confused, but it is often the piecemeal nature of the responses required which is attacked by teachers, whether these responses are creative or otherwise.

The scheme to be described below assumes that many teachers wish for the restoration of the essay as part of the examination pattern. In brief, the solution rests on the combining of assessments of the children's essay writing by their own teachers, and scaling against an external examination. It was thought vital to have this scaling against an essay-writing examination rather than, say, against objective English tests if any reduction of backwash was to be hoped for. The main interest of the paper (apart from the demonstration of the validity of the procedures adopted) is in the re-examination of the rationale of marking and in the detail of the practical arrangements.

The traditional procedure assumes that the marker can allocate marks by some standard or set of standards which are not influenced by the objects of the judgments, but it is a commonplace of experimental psychology that personal-response-scales " readily shift, expand and contract with a change in the context in which the judgments are made " (Postman and Egan, 1949). Wiseman's *Instructions for Marking Essays* (1949) show recognition of ways in which subjective scales can be distorted by irrelevancies, and markers are specifically warned against them. An example of the way in which subjective scales drift or become re-anchored is provided by the Yates and Pidgeon enquiry (1957a and 1958b). " The results also indicate that the standards adopted by the examiners were affected by the pupils' improved performances so that when confronted a second time with scripts that they had marked earlier, they betrayed a tendency to apply higher standards and consequently to award lower scores." Such effects may be expected, not only between marking occasions but also within a marking batch of any length. A run of good papers will tend to make the examiner under-rate the moderate paper which might well have been over-rated if it had been preceded by a run of poor papers. The procedures to be described provide constant checks on the wavering standards of markers.

There are other considerations, some of which seem to ask for no more than a commonsense solution. For example, Finlayson (1951) has shown (and more evidence is offered below) that the reactions of markers to essay titles are idiosyncractic and it would seem an obvious measure to prevent this source of under or over-estimation of essays by judging each title-batch separately.

All the difficulties are not, of course, confined to marking. Vernon and Millican (1954) report a " serious source of inconsistency in assessing English ability in the varying performance of candidates when writing essays on different topics." They were speaking of training college students, but similar causes of misjudgment operate at 11+. This study seems to show that children not only differ from title to title but from type to type of essay so that, unless there is variety in the type of title set, some children may be consistently over- or under-valued.

It is these complexities that the following scheme attempts to meet in an administratively acceptable fashion.

II.—THE PROCEDURES.

(*a*) Primary heads submit an order-of-merit in essay writing. It is the accuracy of the teachers' assessments which is ultimately relied upon and primary heads should be given an outline of the scheme and detailed advice about their part in it. During the six months before the classification-examination, each primary head sets at least three essays of widely different type. Examination numbers are used to preserve anonymity. Each of these essays is marked independently by three members of staff, so that at least nine judgments can be used in making up final orders-of-merit, one for boys and one for girls. The suggested criteria are those used by the external markers. The median judgment is taken as representative in making up the final order. No allowance for age is made on this list.

(*b*) The classification-examinations include an essay. The school lists give the relative merits of children within schools, but do not allow comparisons to be made between children of different schools: therefore, an external examination is also necessary to adjust standards. There is a choice from four titles, which are not restricted to imaginative topics. 30 minutes is allowed.

(*c*) The examination essays are divided into boys' essays and girls' essays, are separated by titles, and are then further sub-divided into suitable marking batches of 60 to 100 papers each. These limits are rough ones based on a compromise between ease of manipulation and the size of the sampling errors introduced. All the boys' essays on one title are place in one pile, the total number is estimated and the pile accordingly distributed to a calculated number of batches, one by one. Such a procedure makes certain that the essays from any school are divided evenly (and at random) between the marking batches. These divisions of the essays can be completed very quickly if the supervisors of the essay-examination are asked to collect the papers so that boys' and girls' essays (and titles) are kept separate in the first place.

(*d*) The marker places the essays of each batch in an order-of-merit, using an agreed basis of judgment described below. Each batch of papers is placed in a provisional order-of-merit, and then this order is altered as necessary after comparison of adjacent papers. Markers are asked to avoid ties if possible.

(*e*) Marking is by the impression method advocated by Wiseman but the markers have agreed on a change of emphasis. The following factors affect placing:

(i) Factors entering into language control, such as vocabulary, variety of sentence structure, clarity and organisation of ideas.
(ii) Wealth of ideas, originality and imagination.
(iii) Spelling, punctuation and grammatical usage.
(iv) Legibility of handwriting and general appearance.

It was agreed that the distinctive contribution of the essay to the existing pattern of objective testing should lie in emphasising (i) particularly, though some variation would arise according to the nature of the titles. As the essays are divided for marking on a title basis such variations of emphasis are possible though at all times the most general criterion, success in communication, is kept in mind.

(*f*) Having arrived at his final order-of-merit, the marker then consults the tables provided which give scores for all possible batches between 21 and 130. The marker finds the column headed by the number in the batch and allocates the scores to the essays in turn. This ensures that the marks from each batch have a mean of 100 and a standard deviation of 15; and the need for checks on the varying standards of markers is thus rendered unnecessary.

(*g*) The marked scripts are then re-ordered by examination-number and the marks entered on to the school examination lists. The marks are now re-distributed according to the orders-of-merit provided by the primary heads. The highest mark is given to the first on his list, the second highest mark to the second child on his list, and so on—an adjustment for age being made at the same time.

(*h*) The application of an age-allowance is quite simple. The children of average age at the time of the examination are allowed nothing, and the others have marks added or subtracted according to the scale in Table 1.

TABLE 1

AGE ALLOWANCE.

	Months younger than average						Ave. Age	Months older than average					
	6	5	4	3	2	1		1	2	3	4	5	6
Correction for age ..	+3	+3	+2	+2	+1	+1	0	−1	−1	−2	−2	−3	−3

If, for example, the average age is 10 : 11, a child of 11 : 2 has 2 subtracted from the score which is transferred to him and a child of 10 : 5 has 3 added to the score transferred to him. The amount of age-allowance has been arrived at empirically, but it may be relied upon to change little unless the standard deviation of the scores is changed, and this is a further advantage of the use of the tables. This procedure for age-allowances may appear as a drastic over-simplification of the usual table provided by Moray House and the N F E R for use with objective tests by the inaccuracy involved is negligible.

(*i*) No special examination is needed for absentees, since any absentee can be given a mark lying between those earned by his neighbours in the head's order-of-merit. If an absentee is at the top or bottom of a head's order-of-merit, it is best to give a mark equivalent to the objective English test mark. These adjustments for absentees have to be made before the re-distribution of marks.

III.—COMMENTS ON THE PROCEDURES.

It is convenient to comment at this point on certain aspects of the scheme.

(*a*) *Time.*

Despite its apparent complications, the total time spent on the essay-examination (administrative, marking and clerical), is less than the time spent on an objective test.

(*b*) *The Tables.*

In the range 21 to 50, the tables are based directly on Table XX of Fisher and Yates (1956). In the range 51–130, they were worked out by the present writer from the normal probability curve. These distributions of scores have standard deviations of slightly less than 15 because of the assumptions made in their construction, but this is of no practical importance. (The writer will send a copy of the tables to any interested reader).

(*c*) *Instructions to Markers.*

Apart from the general instruction to mark on an impression of the whole performance, Wiseman (1949) instructed his markers to act on a basis quite different from the one used in this scheme. They were instructed to look for signs of ability other than literary and to reward the good writing which resulted from intelligence and originality. These instructions appear to be turning the essay into an intelligence test : " You are judging children, not essays." Under this scheme on the other hand, the markers, both school and external, are not expected to make inferences about children but to assess the writing itself as outlined above.

There is, of course, no evidence that instructions to markers affect what they would do on the basis of their previous professional experience. These differences of emphasis seemed advantageous in principle, whatever effects they might have on immediate reliability. It was hoped that the change of emphasis might make the classification-examination as a whole more valid by making the contribution of the essay more individual in character.

(*d*) *Splitting into Batches.*

The case for splitting into batches would be weakened if it could be shown that certain titles were apt to be chosen by the more intelligent children. However, an investigation of the verbal reasoning quotients of six groups choosing different titles showed that the differences of the batch means from the mean of their sex group were not significant. (75 essays per batch.)

(*e*) *Types of Essay.*

Independently of the investigation of the scheme in the classification examination, a subsidiary investigation of marking was undertaken. Sixty-four essays written by sixteen children aged 10 or 11 years were marked (without splitting) by training college students. Each essay was given a different examination mark so that the markers could not know that the essays were not produced by sixty-four different children. The results of this marking investigation confirms those already found : (i) significant differences between titles and between markers ; (ii) significant interactions between children and titles and between markers and titles. This is the evidence which points to the need for multiple assessments of children and division into batches with fixed standards. The investigation did, however, attempt something more than has been reported before.

It was anticipated that descriptive essays would be under-valued by markers and examination of the means of the essays showed that this was so. In terms of scales based on a standard deviatisn of 15 and a mean of 100, a child judged on a descriptive essay was ten marks ' worse ' than when he wrote an imaginative essay.

Table 2 shows the analysis of variance of the scores formed from combining those of two imaginative essays and those of two descriptive essays.

TABLE 2

ANALYSIS OF VARIANCE OF ' SCORES ' FOR TYPES OF ESSAY TITLES.

Source of Variation	d f.	S.S.	M.S.	V.R.
Children	15	55,323·750	3,688·250	—
Markers	3	0	0	—
Type of Essay Title (T E T)	1	12,600·781	12,600·781	—
C×M	45	6,906·250	153·472	—
C×T E T	15	6,032·470	402·167	2·7
M×T E T	3	1,806·595	602·198	4·1
C×M×T E T............	45	6·300·154	140·003	—
Total	127	88,970	—	—

With a revised estimate of error-variance of 146·738, the interactions between types of titles and children, and between types of titles and markers, are both significant at the 1 per cent. level.

This opens out a possible field of investigation but the immediate implications are clear. The external examination ought to encourage all types of writing and the heads should deliberately base their assessments on a wide variety of titles so that a child who is distinctly worse at one type (as some apparently are) then has the opportunity to show merit in other types.

IV.—INVESTIGATIONS AND DISCUSSION.

(*a*) *Reliability of Markers.*

At the first try-out of the scheme, all the papers were marked twice in order to provide data for a reliability check. In view of the high validity-coefficients obtained later, the coefficients for the reliability of markers calculated at this stage are, in a sense, quite unimportant. It may, however, be of interest nevertheless to record that a mean correlation of ·84 was obtained in this study on the basis of re-marking by different markers of 1,279 essays. This compares favourably with the figures obtained by others and usefully brought together by Yates and Pidgeon (1957a, and 1957b) in table form. It is possible that the marking procedures adopted, i.e., order-of-merit marking of one-title batches with fixed standards may have increased the accuracy of judgments.

Consideration of the reliability-coefficient of ·84 leads to examination of the possible gain from using two external markers. It can be shown that the combination of two markers of the above reliability produces a reliability-coefficient of ·91. This decreases the standard error of the 80th percentile by amounts varying from less than one mark for school groups of ten to less than half a mark for school groups of forty or more. That is to say, when scaling is used, there is thus no worth-while practical return from duplicating the work of external marking.

(*b*) *Validity.*

The children whose essay writing was first judged under the new scheme, took their first-year examinations at the grammar schools in the summer of 1961. The three grammar school heads kindly made the results available and thus rendered possible an investigation into the validity of the tests used in the 1960 examination. Choice of the criterion of success was left to the judgment of the heads of the grammar schools., and the criteria used were as follows :

School A (Boys) : The sum of the school examination marks in Mathematics, English and French.

School B (Boys) and School B (Girls) : The sum of English and Mathematics marks.

School C (Girls) : A final mark obtained by adding the marks for English Language, English Literature, French, History, Geography, Scripture, General Science and Mathematics.

The aims of the validity study were to see how the correlations of the allocated essay mark with the criteria compared with the correlations for the objective tests; and also to see how the validity of the total marks formed from the objective tests would be affected by the addition of the essay-mark.

There were two complications. The system of selection, though relying on objective tests, varies in its degree of reliance at the borderzone. At the borderzone an Examination Board procedure places increasing weight upon other factors, especially heads' reports. In addition to this, the final allocation of some children in 1961 was also dependent upon the essay-marks which were available to the Board. These complications might have affected any straightforward comparison between the tests and the essay. Therefore, the children whose total scores were within 15 marks of those of the lowest child admitted were eliminated from the validity study. This reduced the numbers by about one-half. The comparisons based on the higher scoring group may be thought to be biased against the essay correlation, since this group is based on success in the objective tests rather than on a combination of these with the essay.

Because of the difficulties of arriving at age-allowances for each school examination, they were also eliminated from the standard scores given by the objective test tables. The raw scores themselves could not be used because skewing in the distributions distorted the ' boosted ' correlations. The correction formula used was Guilford's (1956) formula 13·29. The average correlations were obtained via z transformations.

TABLE 3

CORRELATIONS OF CLASSIFICATION TESTS WITH THE CRITERIA

(a) Verbal Reasoning (Mean of M H T 62 and 63).
(b) Arithmetic (M H A 30).
(c) English (M H E 30).
(d) Essay Marker (One Marker).
(e) Essay Marker (Two Markers).
(f) Allocated Essay Mark.
(g) Total of (a), (b) and (c).
(h) Total of (a), (b), (c) and (f).

	a	b	c	d	e	f	g	h
School A (Boys)	·89	·60	·87	·76	·82	·87	·91	·97
School B (Boys)	·85	·70	·68	·36	·57	·70	·87	·89
School B (Girls)	·78	·72	·33	·38	·35	·46	·86	·88
School C (Girls)	·93	·79	·87	·02	·22	·63	·95	·93
Average	·87	·71	·74	·42	·54	·70	·90	·93

The main interest of Table 3 lies in the average correlations (based on the results of 121 children). The allocated essay mark has validity of the same order as the arithmetic and English marks. On the other hand, the validity of the external essay marks, even when doubly marked, is distinctly poor. There is no evidence that the validity of the examination based on objective tests is reduced by the addition to the total of the allocated essay marks.

It should be pointed out that the highly satisfactory validity-coefficients may well be due to some ' halo '-effect in the primary heads' orders-of-merit in essay writing. They may import judgments of general suitability based on observation of abilities other than essay writing, unless their school-examinations are conducted with proper regard for anonymity of pupils in the process of marking. This is true, but from a practical point of view, it is unimportant. The administrator is concerned with introducing a valid examination which encourages essay writing within the primary schools.

V.—CONCLUSIONS.

The most important conclusion reached is, of course, that the processes and procedures described do lead to a valid examination. There is no claim that an examination which already consists of, say, three objective tests, will achieve increased validity through the addition of an essay examination ; but the size of the validity-quotient of the essay examination does preclude any possibility of a decrease in the validity of the whole.

Clearly any examination must be considered from an administrative point of view. The following advantages are claimed :

(*a*) There is only one external examination.

(*b*) The procedures can be adapted to cover, by commonsense use of the tables, the whole range of ability or smaller proportions as required: for instance, if the upper 50 per cent. of the school population is being examined, then batches of sixty would have to be marked by reference to the part of the tables headed by 120.

(*c*) The clerical work, entering scores on lists, consulting conversion tables, etc., is no more difficult or lengthy than that for any objective examination, and the statistical work is very much reduced in comparison with that of other possible essay-examinations, in which checks on means, distributions and age-allowances have to be made.

(*d*) There need be no overloading of a few markers. The papers can be split very widely, e.g., 2,000 papers might well be split amongst 30 markers.

(*e*) The system allows great variety in the choice of essays for the external examination as there need be no guessing about the relative difficulty of suggested titles. This is particularly important in view of the apparent need to encourage teachers to make assessments of children based on different types of writing.

(*f*) The scheme is attractive to markers because it relieves them of the difficult task of maintaining constant standards. The time demanded for building orders-of-merit is no greater than that needed for the usual ' numerical ' marking.

VI.—REFERENCES.

FINLAYSON, D. S. (1951). The reliability of marking essays. *Brit. J. Educ. Psychol.*, 21, 126-134.

FISHER, R. A., and YATES, F. (1956). *Statistical Tables*, 5th Ed. Oliver and Boyd.

GUILFORD, J. P. (1956). *Fundamental Statistics in Psychology and Education.* 3rd ed. McGraw Hill.

PIDGEON AND YATES (1957b). Symposium : the use of essays in selection at 11+. IV—Experimental inquiries into the use of essay-type English papers. *Brit. J. Educ. Psychol.*, 27, 37-61.

VERNON, P. E., and MILLICAN, G. D. (1954). A further study of the reliability of English essays *Brit. J. Stat. Psychol.*, 7, 65-74.

VERNON, P. E. (1957). *Secondary School Selection.* Methuen.

WISEMAN, S. (1949). The marking of English composition in grammar school selection *Brit. J. Educ. Psychol.*, 19, 200-209.

YATES AND PIDGEON (1957a). *Admission to Grammar Schools.* Newnes.

(*Manuscript received, 4th March*, 1962.)

SOURCES OF DISSATISFACTION AMONG A GROUP OF TEACHERS

By W. G. A. RUDD and S. WISEMAN
(School of Education, University of Manchester)

SUMMARY. 1.—590 teachers, forming about 72 per cent. of the total output of trained teachers in 1955, from eleven constituent colleges and departments of the University of Manchester School of Education, completed in 1960 a questionnaire reporting, *inter alia*, current levels of satisfaction and major sources of dissatisfaction in the profession.

2.—Results showed that the great majority of subjects, including many then out of teaching, had experienced in the profession a high measure of satisfaction. Those in infant and grammar schools (particularly men teachers in the latter type of school) appeared to derive more satisfaction from their work than did subjects in some other types of school.

3.—Some evidence is presented which suggests that head teachers may be very well satisfied with the levels of professional competence of these subjects; but full analysis of these data is not yet completed. Low positive correlations are reported between the ex-student's practical teaching mark, degree of professional satisfaction and head teacher's rating.

4.—Sources of dissatisfaction appeared to be related more closely to factors in the teaching than to those in the training situation.

5.—Major areas of dissatisfaction included salaries, poor human relations among the staff, inadequate buildings and equipment, high teaching load, training inadequacies, large classes, expressions of personal inadequacy, lack of time for certain professional duties and low status of the profession in society. In listing items women teachers seemed to be more pre-occupied with day-to-day classroom problems, whilst men found their frustrations in a wider context.

I.—INTRODUCTION.

THE data reported in this paper form part of a research being carried out in the University of Manchester School of Education, the major aim of which is to investigate the relationship between qualifications of students on entry to training institutions in the area, and performance during training on the one hand and subsequent success in the teaching profession on the other.

The subjects of this enquiry formed the total output of trained teachers from all constituent colleges and departments of the School of Education who completed their training in 1955. They came from eleven institutions: six general training colleges, a college of housecraft, a training college for technical teachers, a college of art, the University Department of Education and the University Department of Education of the Deaf. The output of these training centres in 1955 was 304 men and 511 women, making a total of 815 subjects for the enquiry. This study is based on the results of a questionnaire sent to each subject in 1960 through the parent colleges and departments.

As might be expected, it proved impossible to trace some of the subjects. However, assiduous enquiry and the co-operation of the Ministry of Education and of appropriate local education authorities finally resulted in replies being received from 590 subjects (72·4 per cent. of total group). One suspects that not all of those who failed to return a questionnaire did so because it never reached them; and it is important to try to separate the truly missing sheep from the wilfully non-co-operative goats, since this vitally affects our ability to

TABLE 1

PERCENTAGE FREQUENCIES.

	Principles of Education		Practical Teaching	
	No response	Response	No response	Response
A	1·1	3·6	2·2	6·4
B	20·1	28·6	29·6	33·1
C	62·6	56·4	58·7	50·9
D	15·6	10·7	7·8	9·0
E	0·6	0·6	1·7	0·6
Total....	100·0	99·9	100·0	100·0
χ^2	$P = \cdot 20$		$P = \cdot 10$	

generalise from the responses received. Table 1 records the distributions of marks for professional subjects in the final examinations obtained by respondents and non-respondents. Chi-squared calculated on the full mark distribution using plus and minus marks from A to E+ showed that differences were not statistically significant. This suggests that the respondents were not unrepresentative of the whole group of subjects on the criterion of performance, whilst students, in these two important tests.

II.—DEGREES OF SATISFACTION.

In the section of the questionnaire under review, subjects were first asked to estimate their current level of professional satisfaction, the actual wording of the item being " How satisfying do you find your job as a teacher ? Try to be as honest and frank as you can in your response. Put a cross in one of the squares." The squares were entitled *fully satisfying, satisfying on the whole but not fully so, moderately satisfying, some satisfaction but not a great deal* and *most unsatisfying*. 539 subjects responded to this item, and their responses are recorded in Tables 2, 3 and 4, as ratings scored from 5 (fully satisfying) to 1 (most unsatisfying).

Table 2 analyses these ratings for all respondents. An important feature of this table is the distribution of totals for all subjects, which shows a strong negative skew, 91·7 per cent. of all responses falling within the fully satisfying or satisfying on the whole categories. Analysis of the totals for subjects still teaching as against those for subjects who had left the profession reveals no significant difference of distribution. This is to be expected, since the overwhelming majority of those not then teaching were in this category for matrimonial reasons and not because of dissatisfaction with, or lack of success in, the teaching profession. Nor was a significant difference found between the distributions of ratings given by men and women teachers still in the profession. From these data, therefore, we conclude that the vast majority of this group of ex-students, both men and women, have found in teaching a high measure of professional satisfaction. Since many of these respondents served less than two years as teachers, much of this satisfaction must have stemmed from the time of their earliest experiences in the profession.

TABLE 2

DEGREES OF SATISFACTION.

ANALYSIS FOR ALL RESPONDENTS BY SEX.

Degrees of satisfaction	Totals						
	S			N.N.S.			T
	M	W	T	M	W	T	
5	81	87	168	2	42	44	212
4	113	111	224	—	58	58	282
3	11	17	28	—	7	7	35
2	3	3	6	—	3	3	9
1	—	—	—	—	1	1	1
W.H.	2	8	10	4	37	41	51
Totals	210	226	436	6	148	154	590

Legend :

S—Teachers in service.	W—Women.
NNS—Teachers not then in service.	T—Total.
M—Men.	WH—Information withheld.

TABLE 3

DEGREES OF SATISFACTION.

ANALYSIS FOR ALL RESPONDENTS BY TRAINING INSTITUTION.

Degrees of satisfaction	Training Institutions											T
	A	B	C	D	E	F	G	H	I	J	K	
5	10	27	3	24	35	2	10	24	30	25	22	212
4	10	28	6	38	54	13	20	29	32	28	24	282
3	1	5	—	2	5	—	5	2	3	4	8	35
2	1	1	—	1	2	—	2	1	—	—	1	9
1	—	—	—	—	—	—	1	—	—	—	—	1
W.H.	6	1	1	7	10	—	8	—	4	14	—	51
Totals	28	62	10	72	106	15	46	56	69	71	55	590

Legend :—As for Table 2.

Table 3 analyses the same data by training institution, and reveals that students from College G rated their level of satisfaction significantly below that of subjects from Colleges I†, J†, B*, D* and E* ; but these differences were not supported by longer lists of sources of dissatisfaction. (See below.) All subjects from College G are, or have been, teachers of housecraft, whereas those in groups I, J, B, D and E came from three general colleges, the training college for technical teachers and the University Department of Education. Had the low levels of satisfaction expressed by these housecraft teachers been due to an accumulation of irritants in their teaching situations, one would expect this to

---

† Difference significant at 1 per cent. level.
* Difference significant at 5 per cent. level.

TABLE 4

DEGREES OF SATISFACTION.

ANALYSIS FOR FULL-TIME TEACHERS SERVING G.B. AND N. IRELAND, BY SEX AND TYPE OF SCHOOL IN WHICH THEN TEACHING.

Degrees of satisfaction	Infant			Junior			Sec. Mod.			Sec. Gram.			F.E.			Special			Total†		
	M	W	T	M	W	T	M	W	T	M	W	T	M	W	T	M	W	T	M	W	T
5	—	26	26	8	19	27	21	23	44	23	6	29	21	2	23	—	3	3	78	85	163
4	—	22	22	17	29	46	37	24	61	27	13	40	22	2	24	5	4	9	115	101	216
3	—	2	2	2	5	7	2	8	10	1	—	1	3	1	4	1	—	1	10	16	26
2	—	—	—	1	1	2	—	2	2	—	—	—	1	—	1	—	—	—	2	3	5
1	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
W.H.	—	2	2	—	—	—	—	2	2	—	1	1	—	—	—	—	—	—	1	5	6
Totals	—	52	52	28	54	82	60	59	119	51	20	71	47	5	52	6	7	13	206	210	416

Legend : As for Table 2.

† Also included in these totals are data for teachers serving in other types of school, as follows : secondary technical (2 men and 2 women each scored 5, 5 men each scored 4 and one man scored 3) ; independent (2 men and 3 women each scored 5, 7 women each scored 4) and training colleges (1 man and 1 woman each scored 5 and 2 men each scored 4, one man withheld information).

be reflected in higher-than-average totals of reported sources of dissatisfaction. Since this was not so, and all respondents then teaching were serving in secondary schools, we can only speculate that the teaching situation in such schools permits, perhaps, fewer *positive* satisfactions for housecraft teachers than for those serving in other circumstances.

Table 4 includes only data from subjects in full-time service as teachers in Great Britain and Northern Ireland. As before, there are no significant differences in the distributions of total ratings given by men and women teachers. However, when these totals are broken down by type of school, ratings from infant teachers emerge as significantly higher than those from junior and secondary modern teachers,* whilst grammar school teachers rate their satisfaction significantly higher than do their counterparts in the junior schools.* Wherever in this table totals were sufficiently large, comparisons were made between the following rating distributions: men still teaching *v.* women still teaching; subjects teaching in schools of one type *v.* those in schools of every other type; men *v.* women teachers in schools of the same and different types. Only one significant difference was found, that men teachers in grammar schools rated their levels of satisfaction higher than did men teachers in junior schools.* Since these findings are supported by others to be reported in this article, further discussion of these data will be deferred until later.

Elsewhere in the questionnaire permission was sought from each teacher to ask for a confidential report of the kind supplied to support an application for another post; and 417 (70·7 per cent.) of the 590 respondents gave this permission. In addition to writing this report, each head gave an overall rating of the teacher in terms of the following six-point scale: excellent, 6; very good, 5; good, 4; good average, 3; average, 2; rather weak, 1. Analyses of these ratings, received in respect of 282 of the subjects are given in Tables 5, 6 and 7.

Table 5 analyses these data for all respondents by sex; and, as before, there is a strong negative skew in the distribution of total ratings, 80·4 per cent. of all responses falling in the top three categories of the scale. This suggests either that head teachers are very well satisfied with the competence of these

TABLE 5

RATINGS BY HEAD TEACHERS.

ANALYSIS FOR ALL RESPONDENTS BY SEX.

Head's rating	Totals						
	S			N.N.S.			T
	M	W	T	M	W	T	
6	34	31	65	1	3	4	69
5	42	47	89	—	10	10	99
4	35	28	63	—	6	6	69
3	10	8	18	—	5	5	23
2	5	6	11	—	5	5	16
1	3	2	5	—	1	1	6
Totals	129	122	251	1	30	31	282

Legend: As for Table 2.

members of their staffs, or that their charitable instincts have come into play with some force. The former of these alternatives may be more likely, since no significant difference was found in the distribution of ratings given for women not then teaching as against that for women still teaching. However, the references themselves are being rated and also subjected to a content analysis. The results of this further enquiry (not yet completed) may throw further light on this issue.

TABLE 6

RATINGS BY HEAD TEACHERS.

ANALYSIS FOR ALL RESPONDENTS BY TRAINING INSTITUTION.

Head's rating	Training Institutions											T
	A	B	C	D	E	F	G	H	I	J	K	
6	4	12	—	9	11	—	4	8	9	2	10	69
5	4	16	—	7	21	—	6	10	18	6	11	99
4	1	8	—	8	21	—	—	8	7	9	7	69
3	—	3	—	1	4	—	4	2	4	4	1	23
2	—	—	—	1	1	—	2	3	4	4	1	16
1	—	2	—	1	1	—	1	—	—	1	—	6
Totals	9	41	—	27	59	—	17	31	42	26	30	282

Legend : As for Table 2.

Comparison of the data for individual training institutions (in Table 6) shows that ex-students of College J scored ratings significantly lower than those from Colleges B†, E†, H†, K†, D* and I*. It would be surprising if differences of this kind did not reveal themselves in an inter-college analysis. The interpretation of them, however, is not so easy. The temptation to assume that they point to significant differences in the quality of output caused by poorer training methods must be resisted until further data are collected, including, in particular, information on the quality of student entrants. The criterion itself is far from satisfactory and must be buttressed by the results of the fuller analysis of the references themselves (referred to above) and related to independent criterion measures derived from biographical data. The possibility of inter-action with other factors must also be taken into consideration.

Table 7 groups the ratings given for subjects teaching in schools of different types in Great Britain and Northern Ireland. Wherever totals are sufficiently large the comparisons listed earlier for data in Table 4 were made also for these data. In only one instance did a significant difference emerge, that subjects teaching in further education establishments gained ratings significantly higher than those gained by subjects teaching in secondary modern schools.*

Next, product-moment correlations were computed between the subject's expressed degree of satisfaction, his practical teaching mark gained whilst a

---

† Difference significant at 1 per cent. level.

* Difference significant at 5 per cent. level.

TABLE 7

RATINGS BY HEAD TEACHERS.

ANALYSIS FOR FULL-TIME TEACHERS, BY SEX AND TYPE OF SCHOOL IN WHICH THEN TEACHING.

Head's rating	Infant			Junior			Sec.Mod.			Sec. Tech.			Sec. Gram.			F.E.			Total†		
	M	W	T	M	W	T	M	W	T	M	W	T	M	W	T	M	W	T	M	W	T
6	—	8	8	4	10	14	10	8	18	2	1	3	6	1	7	11	1	12	34	29	63
5	—	7	7	5	9	14	8	15	23	1	2	3	12	8	20	15	1	16	41	43	84
4	—	8	8	2	6	8	13	9	22	2	—	2	11	2	13	5	—	5	34	26	60
3	—	1	1	2	2	4	2	3	5	—	—	—	5	—	5	1	—	1	10	6	16
2	—	2	2	—	1	1	4	2	6	—	—	—	1	—	1	—	1	1	5	6	11
1	—	1	1	1	1	2	1	—	1	—	—	—	—	—	—	1	—	1	3	2	5
Totals	—	27	27	14	29	43	38	37	75	5	3	8	35	11	46	33	3	36	127	112	239

Legend : As for Table 2.

† Included in these totals are scores for 2 men teaching in training colleges (one rated 6 and the other 4) and for 2 women teaching in special schools (one rated 5 and the other 4).

TABLE 8

PRODUCT-MOMENT CORRELATIONS.

	Practical Teaching Mark			Head Teacher's Rating		
	Men	Women	Total	Men	Women	Total
Degree of satisfaction	·106 (N=208)	·008 (N=213)	·051 (N=421)	·234 (N=129)	·159 (N=122)	·208 (N=251)
Head Teachers' rating	·235 (N=129)	·138 (N=122)	·187 (N=251)	—	—	—

student and the head teacher's ratings. Table 8 records the results of these computations. It will be seen from this table that in all cases a slight and positive relationship exists between these ratings and the degree of satisfaction expressed by the teacher. In both cases that relationship is higher for men than for women subjects.

III.—SOURCES OF DISSATISFACTION.

In addition to estimating his present level of professional satisfaction, each subject was asked in the questionnaire to list his chief sources of professional dissatisfaction (if any). Elsewhere, the questionnaire also invited the respondent to add any other information considered relevant to his development as a teacher since leaving college; and with the questionnaire went a letter from a college tutor inviting the ex-student, if he so wished, to communicate more personal data in a covering letter when returning the questionnaire. All this produced a large and heterogeneous mass of data which could be handled only by some system of classification; and this classification was built up empirically.

First, the replies received were scanned to discover certain broad areas of dissatisfaction. Next, all data were recorded, and these areas divided into categories as recording proceeded. During this process additional areas emerged; and these also were categorised in similar fashion. Since great care was taken not to combine data implying even a moderate difference in meaning, this process resulted in a list of 182 categories. By review and re-grouping the list was cut down to 91 categories, which appeared to the investigators to be disparate. All data were then recorded according to the revised classification. Still later, the 91 categories were combined into 19 major categories; and it is on these last that most of the ensuing analysis is based. However, to avoid the danger of our data merely demonstrating what our categorisation has assumed, reference will be made to the evidence of the 91 categories for all aspects of the teaching situation where dissatisfaction is widespread.

The major categories of sources of dissatisfaction are as follows:

A. Pupils' behaviour (e.g., lack of apparent ability).
B. Pupils' attitudes (e.g., lack of respect for teachers).
C. Teacher training (criticisms of training course).
D. Deficiencies of teaching skill (e.g., difficulties with discipline).
E. Feelings of inadequacy (e.g., self-deprecation).
F. Teaching load (e.g., extra-curricular activities, school meals).
G. More time needed (e.g., for individual pupils; for study and research).
H. School organisation too rigid (e.g., approach too academic).
I. School organisation too loose (e.g., insufficient attention to basic subjects).
J. Other criticisms of school programme (e.g., lack of aims and sense of purpose).
K. School buildings and equipment (e.g., overcrowding, shortage of materials).
L. Interruptions to lessons.
M. Noise.
N. Poor human relations among staff (including head, inspectors and administrators).
O. Poor relations with parents.
P. Poor relations with industry.
R. Salaries.
S. Status of the profession in society.
X. Large classes.

Table 9 analyses, by training institution, the replies of all respondents to the questionnaire. Between them the 590 subjects produced 814 separate sources of dissatisfaction, an average of 1·38 items per respondent. This low average supports the finding reported earlier that this group of teachers is well

TABLE 9

Sources of Dissatisfaction.

Average Numbers of Items Reported by All Respondents Analysed by Sex of Teacher and Training Institution.

N	Training Institution	Average No. of items reported	Average, excluding category R
28	A	0·68	0·64
62	B	1·18	1·03
10	C	0·50	0·50
72	D	1·19	1·10
106	E	1·40	1·23
15	F	2·00	1·87
46	G	1·06	1·06
56	H	2·20	1·41
69	I	1·42	1·36
71	J	1·00	0·99
55	K	1·46	1·38
216	Total men	1·69	1·40
374	Total women	1·20	1·17
226	Total women then teaching ..	1·23	1·18
148	Total women not then teaching	1·16	1·15
590	Total	1·38	1·25

satisfied with its experiences in the profession ; and it stands in marked contrast to the image of the teaching profession as a body of militant malcontents frequently built up in the popular press when new salary scales are under consideration. Men had more grumbles than women (1·69 as against 1·20), while women then in the profession averaged 1·23 grumbles as against 1·16 for those who had left the profession. No doubt the slight difference here represents the emollient effect of the passage of time. On the other hand, it might be argued that there would tend to be a negative correlation between marriageability and tendency to disgruntlement, thus producing the situation of the more satisfied women teachers having a higher marriage rate than the remainder. In any case the data, together with those presented in Table 2 suggest strongly that dissatisfaction with their experience in the profession was unimportant as a reason for these women, as a group, leaving the profession.

When analysed by training institution averages varied between 0·50 and 2·20. Clearly, no conclusions can be drawn in respect of institutions whose ex-students recorded fewer-than-average sources of dissatisfaction. In the cases of two institutions returning substantially more than average scores, the data indicate that factors in the teaching, rather than in the training, situation underlie the scores. Respondents for institution H are all men, teaching almost without exception in junior and secondary modern schools; and we have shown above that such subjects tend to be less satisfied than do other groups. (Table 4.) Institution F prepares teachers for highly specialised work. Of the thirty items listed for this group of subjects, nine are in major category K (shortage of teaching equipment, inadequate accommodation, overcrowded classrooms), and a further six in major category N (lack of support from head teacher, dislike of attitudes of other members of the staff). Apart from the two checks each concerning salaries and the training course, the remaining items fill out the picture suggested by these major sources of dissatisfaction, and include: large classes, not teaching the subjects for which the respondent was trained, more effective streaming wanted, more guidance wanted from the head, parents expect too much from the teacher and married woman unable to give of her best at school and at home.

TABLE 10

SOURCES OF DISSATISFACTION.

AVERAGE NUMBERS OF ITEMS REPORTED BY 416 RESPONDENTS THEN TEACHING IN SCHOOLS OF DIFFERENT TYPES IN GREAT BRITAIN AND N. IRELAND.

N	Type of School*	Average numbers of items reported	Averages excluding category R
28	Junior, men	2·57	2·04
60	Second modern, men	1·90	1·53
59	Secondary modern, women	1·49	1·44
20	Secondary grammar, women	1·45	1·40
47	Technical and F.E., men	1·36	1·21
54	Junior, women	1·35	1·28
52	Infant, women	1·15	1·13
51	Secondary grammar, men	1·00	0·84
10	Secondary technical, total	2·50	2·00
82	Junior, total	1·77	1·54
119	Secondary modern, total	1·70	1·49
13	Special, total	1·46	1·23
52	Technical and F.E., total	1·37	1·23
52	Infant, total	1·15	1·13
71	Grammar, total	1·11	1·00
12	Independent	0·42	0·42
416	All respondents, total	1·47	1·31

* Data omitted for groups of fewer than ten respondents.

Table 10 records average numbers of items reported by respondents then teaching in different types of school in Great Britain and Northern Ireland. These data confirm the evidence presented in Table 4, in that the infant and grammar schools (the latter particularly for men teachers) stand out clearly as the most satisfying types of school for this group of teachers. It is arguable that such a result merely reflects the profession's concern with salaries. For the most part salary levels are a grumble of men teachers; and there are no men teachers in infant schools, whilst many of those in grammar schools and further education establishments receive higher-than-basic salaries. However, our data refute this argument. The last column of Table 10 excludes from the averages all grumbles about salaries; and the rank order of the schools remains virtually the same in this as in the previous column.

It is interesting to compare the rank orders for the two sexes. If we exclude technical and further education (too few women teachers represented) we get almost a complete reversal of order for women as compared with men teachers.

The women find less and less cause for dissatisfaction as the age of the pupils decreases: for men, exactly the opposite occurs.

It is interesting to speculate on possible reasons for the high degree of satisfaction expressed by infant teachers. This group is outstanding for the low scores recorded in major category N (poor human relations in school), and it may be that this, coupled with the satisfactions of the teacher-pupil relationship in such schools, more than offsets the difficulties of organisation brought about by large classes. Nevertheless, of the checks given by infant teachers, 20·0 per cent. went to large classes, 11·7 per cent. to criticisms of the training course,* and 10.0 per cent. each to more time needed, school organisation too loose, inadequate school buildings and equipment. Clearly, the major problems facing this group of teachers centre on the organisation of their classes. In spite of this the teachers derive a high degree of satisfaction from their daily work, no doubt because human relations seem to cause less strain for women in schools of this type than elsewhere. By contrast, grammar school teachers recorded 19·0 per cent. of their checks in major category N, 11·4 per cent. to feelings of inadequacy and 10·1 per cent. each to teaching load and salaries. When the further points are made that three of the nine checks recorded under feelings of inadequacy refer to anxiety about promotion prospects and a further four to a desire for further courses of training (possibly to improve those prospects) some indication is given of the tensions experienced by grammar school teachers in this stage of their professional careers.

Data in Table 10 also confirm earlier findings concerning the relatively low levels of satisfaction experienced by subjects teaching in junior and secondary modern schools, particularly by men teachers. (Very few subjects replied from secondary technical schools, but those who did had a substantially higher average of grumbles than had their counterparts in secondary modern schools). Dissatisfaction with salary scales figures very prominently among the grumbles of men teachers in both junior and secondary modern schools; but Table 10 makes clear that, even when those are excluded, these groups of teachers continue to return well-above-average scores. Within their own totals junior school teachers give 16·5 per cent. of their checks to large classes (men 9·7 per cent., women 23·3 per cent.), 13·1 per cent. to inadequate school buildings and equipment (men 12·5 per cent., women 13·7 per cent.), and 13·1 per cent. to salaries (men 20·8 per cent., women 5·5 per cent.). Among other major categories substantial differences exist between men and women junior school teachers in respect of poor human relations (men 11·1 per cent., women 1.4 per cent.), and as regards the status of the profession in society (men 9·7 per cent., women 1·4 per cent.). These data suggest that the relatively high degree of dissatisfaction expressed by this group of junior school teachers has complex origins and that the grumbles recorded may well be symptoms of a more fundamental malaise.

A rather different picture emerges from the data provided by subjects teaching in secondary modern schools. In line with their counterparts in junior schools men and women differed in the percentage of their total scores given to salaries (total 12·4 per cent., men 19·3 per cent., women 3.·4 per cent.) and to professional status (total 5·4 per cent., men 7·9 per cent., women 2·4 per cent.). Elsewhere, however, much more consensus exists between the sexes in modern than in junior schools, for example towards poor human relations (total 11·9 per cent., men 10·6 per cent., women 13·6 per cent.); inadequate buildings and equipment (total 9·4 per cent., men 8·7 per cent., women 10·2 per cent.); criticisms of the training course* (total 9.9 per cent., men 10.5 per cent., women 9·2 per cent.). It would thus appear likely that different causes underlie the relatively high levels of dissatisfaction expressed by these groups of junior and secondary modern school teachers. If it is assumed that public opinion will accept only a limited increase in expenditure on existing education services, then, to the extent that these data represent widely-held attitudes among teachers, it is clear that neither reducing the sizes of existing classes, nor raising teachers' salaries, nor improving existing standards of school accommodation nor any other of the universal panaceas currently being advocated by interested parties is likely, of itself, to banish feelings of dissatisfaction among teachers in junior and secondary modern schools.

---

* To some extent data in this major category are inflated, since the letter accompanying the questionnaire invited specific comments on the adequacy of the training course.

TABLE 11

SOURCES OF DISSATISFACTION.

RANK ORDER OF MAJOR CATEGORIES FOR 416 FULL-TIME TEACHERS.

Rank order	Category	Frequency of mention
1	R. Teachers' salaries	68
2	N. Poor human relations among staff	66
3	K. Inadequacies of school buildings and equipment	63
4	F. Teaching load	57
5½	C. Teacher training	50
5½	X. Large classes	50
7	E. Feelings of inadequacy as a teacher	42
8	G. More time needed	33
9	S. Status of the profession in society	32

Turning next to the areas of dissatisfaction themselves, Table 11 records in rank order major categories to which subjects then in full-time teaching service allocated 5 per cent. or more of their total checks. Consideration will now be given to the content of these categories.

(1) *R. Teachers' Salaries.*
Four subdivisions are included under this heading:

R1. Salaries too low; maximum reached too slowly; pension rights inadequate; insufficient consideration for industrial experience in fixing salaries. (Men 46, women 7 checks).

R2. Lack of remuneration for extra duties in school. (Men 6, women nil checks).

R3. Problems concerned with the 'rat race' for promotion. A check here implied major concern with the financial aspects of the problem. If the major concern was with the effects on human relations, the check was given under category N4. (Men 4, women 2 checks.)

R4. Anomalies created by the present system for distribution of special responsibility grants; great differences found in salary scales in respect of teachers with similar qualifications. (Men 2, women 1 check.)

(2) *N. Poor human relations among the staff (including head, inspectors and administrators).*

This major category comprised six divisions :

N1. Lack of interest or initiative among staff ; no co-operation among staff ; dislike of other members of staff ; lack of esteem among staff for my subject; criticism of my work or aims by other members of staff. (Men 11, women 4 checks.)

N2. Would like more contact with other members of staff ; not sufficient staff meetings. (Men 5, women 2 checks.)

N3. Lack of co-operation or poor relationship with head (or others in authority, L.E.A., Governors) ; headmaster displays authoritarian attitudes ; lack of support from head in matters of discipline ; lack of consideration from L.E.A. ; inefficiency of educational administration. (Men 18, women 7 checks.)

N4. Human relations problems related to graded posts, shared responsibility, etc. (Men 12, women 4 checks.)

N5. H.M. Inspectors exercise too much influence. (Men 2, women nil checks.)

N6. Lack of sympathy of head of junior school with infant school and its methods. (Men nil, women 1 check.)

(3) *K. Inadequacies of school buildings and equipment.*

Seven sub-divisions were included in this category but only three, accounting for 55 of the 63 checks, will be detailed here :

K1. Overcrowding in school or classroom. (Men 4, women 3 checks.)

K2. Inadequate teaching accommodation. (Men 11, women 14 checks.)

K3. Shortage of books, teaching materials or equipment. (Men 15, women 8 checks.)

(4) *F. Teaching Load (including extra-classroom duties).*

This major category comprised seven divisions, of which four (totalling 49 checks) will be detailed :

F1. Heavy teaching load. (Men 4, women 6 checks.)

F3. Duties other than teaching take up much valuable time. (Men 7, women 6 checks.)

F4. Too much clerical work. (Men 5, women 3 checks.)

F5. School meals supervision. (Men 10, women 8 checks.)

(5½). *C. Teacher training.**

Eleven sub-divisions were included in this category. Four of these, totalling 38 checks will be detailed. None of the remaining categories scored more than 3 checks :

C2. More school practice, observation and/or teaching. (Men 6, women 3 checks.)

C3. More attention in training course to methods of teaching. (Men 9, women 6 checks.)

C8. More clerical training (keeping of school registers, records, etc.) in training course. (Men 2, women 4 checks.)

C11. Regular compulsory leave of absence from schools, for gaining additional experience or training. (Men 4, women 4 checks.)

(5½) *X. Large Classes.*

There were no sub-divisions of this major category.

---

* See footnote to page 286.

(7) *E. Feelings of inadequacy as a teacher.*
This major category included six sub-divisions. Three of these, totalling 35 checks, will be detailed :

E1. Own insecurity ; inexperience ; found teaching very difficult ; personal inadequacy as a classroom teacher. (Men 13, women 10 checks.)
E2. Teaching does not stretch me mentally ; dullness of repetitive work ; boredom of marking. (Men 3 women 2 checks.)
E6. Poor promotion prospects. To merit a check in this category there must be evidence that the subject felt personally inadequate. If the dissatisfaction were externalised, the item would be scored under, for example, categories R3 or N4 (q.v.). (Men 6, women 1 check.)

(8) *G. More time needed.*
Five sub-divisions were included in this category, but only three, accounting for 30 checks, will be detailed here :

G1. For individual work with pupils, e.g., in reading, or those in need of special care and attention. (Men 3 women 12 checks.)
G2. To prepare lessons adequately ; to mark work adequately ; to do work conscientiously. (Men 4, women 5 checks.)
G3. For further study, for individual research, to broaden outlook. (Men 4, women 2 checks.)

(9) *S. Status of the profession in society.*
There were four sub-divisions of this major category :

S1. The low esteem given by the public to education, to particular types of school. (Men 12, women 3 checks.)
S2. Unqualified teachers lower the status of the profession ; insufficient attention given to training when appointments are made. (Men 12, women 1 check.)
S3. Lack of professional unity ; inter-union rivalry. (Men 3, women nil checks.)
S4. Proportion of men to women on mixed staffs is too low. (Men 1, women nil checks.)

Finally, let us consider the rank order of areas of dissatisfaction by sex. Table 12 records this for men and Table 13 for women in full-time teaching at the time of the survey ; and in each case a line has been drawn to separate the major categories (those receiving more than 5 per cent. of the total checks given) from the remainder.

The first difference to be noted is in the spread of dissatisfaction. For women ten areas are included in the upper division of the table, but for men only seven. More significant still, however, are the major categories included in those divisions. For women, but not for men, the following are included : large classes, more time needed, school organisation too rigid, deficiencies of teaching skill and school organisation too loose. Two categories only are included for men but not for women, salaries and the status of the profession in society. This seems to suggest that, to the extent that one can generalise on the basis of this sample, women are more pre-occupied with day-to-day classroom problems and stresses whilst men appear to find their frustrations in a wider context.

One may seek corroboration of this trend among data for the major categories common to the upper divisions for both sexes. Apart from the higher

TABLE 12

Sources of dissatisfaction.

Rank Order of Major Categories for 206 Full-time Men Teachers.

Rank	Category	No of mentions
1	R. Teachers' salaries	58
2	N. Poor human relations among staff	48
3	K. Inadequacies of school buildings and equipment	34
4	F. Teaching load	30
5	C. Teacher training	29
6	S. Status of the profession in society	28
7	E. Feelings of inadequacy as a teacher	24
8	X. Large classes	14
9½	H. School organisation too rigid	13
9½	G. More time needed	13
11	J. Other criticisms of school programme	11
12½	I. School organisation too loose	9
12½	B. Pupils' attitudes	9
14	D. Deficiencies of teaching skill	8
15	O. Poor relations with parents	6
16	A. Pupils' behaviour	5
17½	L. Interruptions to lessons	2
17½	P. Poor relations with industry	2
19	M. Noise	0

TABLE 13

Sources of Dissatisfaction.

Rank Order of Major Categories for 210 Full-time Women Teachers.

Rank	Category	No. of mentions
1	X. Large classes	36
2	K. Inadequacies of school buildings and equipment	29
3	F. Teaching load	27
4	C. Teacher training	21
5	G. More time needed.	20
6½	E. Feelings of inadequacy as a teacher	18
6½	N. Poor human relations among the staff	18
8½	H. School organisation too rigid	16
8½	D. Deficiencies of teaching skill	16
10	I. School organisation too loose	15
11	B. Pupils' attitudes	13
12	J. Other criticisms of school programme	11
13	R. Teachers' salaries	10
14	O. Poor relations with parents	8
15	S. Status of the profession in society	4
16½	L. Interruptions to lessons	2
16½	M. Noise	2
18	A. Pupils' behaviour	1
19	P. Poor relations with industry	0

ranking given by men to problems of human relations, there is little difference in the orders given to the major categories. One may go further and assume that the ratio between the total checks for men and women in any one major category will be roughly equal to that found between the comparable scores for each of its constituent categories. If data for the fourteen common categories listed earlier are scrutinised in this way, anomalies will be found in only five of them. Of these five women teachers overscore on categories K2 (inadequate teaching accommodation) F1 (heavy teaching load) and C8 (wish for more clerical training), whilst men overscore on categories K3 (shortage of books, etc.) and E6 (poor promotion prospects). To some extent these discrepancies support the thesis ; but at this level of analysis score frequencies are really too low to provide firm support for such hypotheses.

IV.—CONCLUSIONS.

During this enquiry we have been particularly impressed by the high levels of professional satisfaction expressed, both positively (degrees of satisfaction) and negatively (low frequencies of sources of dissatisfaction), by our subjects. It appears that these satisfactions are obtainable from the time of a teacher's earliest experiences in the profession. In spite of having to face frustrating classroom conditions infant teachers were particularly satisfied. Grammar school teachers also enjoyed much satisfaction in their work, despite anxiety then being experienced concerning prospects of promotion to more senior posts. Men teachers concentrated many of their major grumbles in such areas as salaries, poor human relations, status of the profession and feelings of personal inadequacy; but very few felt sufficiently strongly about these, or other problems, to leave the profession. By contrast women teachers spread their grumbles over a rather wider field, and appeared to be more pre-occupied with day-to-day classroom problems. Above all, the irritations reported by this group of subjects made clear that feelings of dissatisfaction would not have been banished by increased public expenditure on salaries, buildings or reducing the size of classes. Apparently, however, much benefit could accrue at little, if any, expense through the improvement of human relations in schools.

This analysis has been based on replies from some 70 per cent. of ex-students entering the teaching profession in 1955 from training institutions in the area of the University of Manchester. We believe our findings to be representative of the views of all teachers in that group; but, clearly, we have no way of knowing how typical they are of the views of the profession at large. We suggest that further study in this area might pay particular attention to the frustrations experienced by teachers in junior and secondary modern schools, as well as extending the enquiry to cover older teachers of long standing in the profession.

(*Manuscript received, 11th April, 1962.*)

CRITICAL NOTICE

The Psychology of Creative Ability

GETZELS, J. W., and JACKSON, P. W. (1962). *Creativity and Intelligence.* London and New York : John Wiley, pp. xviii+293, 49s.

BY CYRIL BURT

The interest in what is now commonly termed ' creativity ' goes back to Galton's pioneer studies of the intellectual characteristics of men of genius. Nearly all the eminent men in his list—the scientists, statesmen, and military commanders, quite as much as the poets, painters, or writers of fiction—were distinguished by the originality of their ideas—the " fluency and freedom of their associations " ; and he was particularly anxious to determine in what ways the free and fluent imaginations of the genius differed from those of the insane, the delirious, the drug-addict, and the mere day-dreamer. The earliest attempts to provide an adequate theory of the subject came from the leading critics of the current associationist doctrines—notably Ward and Stout, and were based chiefly on the results of introspective analysis. Stout, in his *Manual,* distinguished two main types of ' ideation'—' serial ' (or ' reintegrative ') and ' divergent ' : only the former—the type of thinking in which the revived experience is reinstated either completely or in part—could be explained by a psychology based solely on the principle of "association by temporal contiguity " (*op. cit.*, Bk, IV, ch. i, ' Trains of Ideas '). A few contemporary associationists—notably Bain—had sought to supplement the traditional theory by recognizing a second principle, namely, " association by similarity." But similarity is not a mode of association ; it is a relation. And Stout's own explanation was a development of Thomas Brown's principle of " relational suggestion," or (as Spearman later termed it) " the eduction of relations and correlates." This he took to be the real key to the creative processes of genius, and illustrated by apt examples from poets like Tennyson and Browning and scientists like Newton.

From an empirical standpoint, however, perhaps the important contribution of the British school was the attempt to attack the problem experimentally by the methods of factor analysis. Galton's original plan for systematic mental testing had included tests for what he called " higher mental processes " ; and several of his earlier disciples endeavoured to construct practicable tests of ' productive thinking ' along lines suggested by these introspective studies. Moore and I, for example (*J. Exp. Pedag.*, I, 1911, pp. 363f), ventured to introduce (among other methods) what we called the ' analogies test ' (based on Stout's theory of relational suggestion), as well as Galton's test of free association (to measure ' fluency '), and Binet's inkblot test (to measure ' apperception '). Further tests were devised by Garnett, Webb, and several of Spearman's research students. Hargreaves in a Monograph Supplement to the *British Journal of Psychology* (1927) published an elaborate research on " The Faculty of Imagination " ; and Spearman summed up his own views in a small volume on *The Creative Mind* (1930).

Most of the British work was carried out with children of school age. In America a number of workers tried somewhat similar procedures to assess creative ability both in younger children and in adults. The second edition of Whipple's *Manual of Mental and Physical Tests* (1915) describes the Analogies test and a new test for Uncontrolled Association ; and his chapter on " Tests of Imagination and Invention " includes the Inkblot test and a test for Linguistic Invention. Chassell devised a number of suggestive " Tests for Originality " to be used with adults (*J. Educ. Psychol.*, VII, 1916, pp. 317f) ; and Simpson in a paper on " Creative Imagination " (*Amer. J. Psychol.*, XXIII, 1922, pp. 234f) sought to demonstrate its relative independence of intelligence. However, the increasing domination of behaviouristic theories diverted attention to those types of mental activity—or (as they preferred to say) of human behaviour—which could be interpreted in terms of the simple principles of mechanical association. Concepts like ' imagination ' or ' productive thinking ' savoured too much of discredited introspectionist doctrines, and were deliberately excluded from behaviourist text-books.

The revival of American interest in ' creativity ' seems to have started with Professor Guilford's presidential address to the American Psychological Association, which dealt specifically with that subject (" Creativity," *Amer. Psychologist*, V, 1950, pp. 444f) ; and most of the newer American work in that field has been inspired, directly or indirectly, by the researches of Guilford and his collaborators at the University of South California. Their work has been concerned almost entirely with adults, particularly with what they call " high level personnel." The same holds true of the inquiries undertaken by Professor Calvin Taylor and his fellow workers at the University of Utah, although their researches have dealt chiefly with 'scientific creativity.' Of the various studies carried out on pupils of school age, the most influential have been those of J. W. Getzels, Professor of Educational Psychology in the University of Chicago. He and his associates have been interested more especially in the creative capacities exhibited by young adolescents. Much of their work has been repeated with younger children by Professor Torrance, Director of Educational Research at the University of Minnesota, who is himself one of the most active investigators in this field.

The volume now published by Professor Getzels and his colleague, Professor Jackson, describes in detail the most important of their various researches. In a prefatory note they relate how, in 1957, Professor Chase, Chairman of the Department of Education in their University, called a meeting of the faculty to discuss the need for research on giftedness. The outcome was the first of the studies reported in this book. It was the work of a co-operative team, financed in part by a grant from the United States Office of Education. The investigators began by asking what particular qualities are generally regarded as most characteristic of the gifted child, and most likely to be prognostic of adult success. As a result of questionnaires addressed to teachers and parents, they concluded that the most important qualities could be classified under four broad headings—intelligence, creativity, emotional adjustment, and moral character. The first pair in their view designate contrasted types of ' cognitive excellence,' and the second pair contrasted types of ' psychosocial excellence.' Accordingly they decided to select four groups of pupils—each pre-eminently distinguished by one or other of these four characteristics, and subject them to a systematic study.

The first and larger part of their volume is devoted to an investigation of the two main types of cognitive activity as thus defined. The hypothesis which they set themselves to verify is that suggested in Guilford's address on " Creativity " : " if the correlations between intelligence test scores and many types of creative processes are only moderate or low—and I predict they will be—it is because the primary abilities represented in the tests are not all-important for creative behaviour, and some of the primary abilities important for creative behaviour are not represented in the tests." To assess intelligence they used the scores obtained with the Binet, WISC or Henmon-Nelson tests. To measure creativity five special tests were devised (based partly on the earlier work of Guilford)—' completing fables,' ' making up problems,' ' hidden shapes,' ' word association,' and ' uses for things.'

The pupils chosen for intensive study were the boys and girls attending the senior classes of a private school in Chicago, ranging from the sixth grade up to the senior year of high school. Those selected comprised (*a*) twenty-eight adolescents falling within the top 20 per cent. in a test of intelligence (average I.Q. 150) but below the top 20 per cent. in tests of creativity, and(*b*) twenty-six falling within the top 20 per cent. as judged by the tests of creativity but below the top 20 per cent. in the test of intelligence (average I.Q. only 127). The greater part of their research consists in the attempt to determine " what the pupils so selected are like (1) as ***students***, (2) as ***individuals***, and (3) as ***members of their family*** " ; and a variety of supplementary tests and inquiries were carried out to discover in what ways, if any, the two contrasted groups differed when studied from each of these three aspects.

In tests of school attainments, the pupils in the ' creative ' group, in spite of their relatively low intelligence, were, if anything, slightly superior to those in the ' intelligent ' group. In tests of ' verbal ' and ' non-verbal imaginative production ' they revealed more humour, playfulness, and wit, and a greater fondness for aggressive and violent situations : most of them were liable to " break away from the stimulus," using it chiefly as a point of departure for free self-expression. In behaviour and in aspiration—e.g., in their choice of careers—they tended to diverge more widely both from the expectations of their teachers and from the patterns dictated by conventional standards. It is, therefore, scarcely surprising that, when the teachers were invited to rate the children for ' desirability,' it was, on the whole, the pupils in the ' intelligent ' group who were assessed as the most desirable members in a classroom. There were marked differences in the home backgrounds. Most of the fathers and more than half of the mothers of the ' intelligent ' group had received a ' graduate training ' : in more than 60 per cent. of the cases the fathers' occupational status is described as " University teaching, research, or editing." Among the ' creative ' group about 50 per cent. of the fathers appear to have been engaged in ' business '—nearly three times as many as in the ' intelligent ' group. The parents of the ' intelligent ' group exercised greater ' critical vigilance ' in regard to their children's day-to-day life and choice of friends, and were able to provide them with a more " bookish ' environment. The main conclusion seems to be that the atittudes both of teachers and of the more academic types of parent tend to favour the educable pupil rather than the creative pupil, and to repress instead of fostering any incipient inclination towards originality or enterprise.

Even so, what the authors consider to be the most puzzling phenomenon of all remains unexplained—the fact that the ' creative ' group, despite their inferior home background, their inferior intelligence, and their inferior behaviour in the classroom, actually score higher marks in the educational tests. This unexpected success, we are told, " places them, as such phenomena are handled to-day, not in the category of ' gifted children,' but in the pejorative category of ' over-achievers '." Accordingly, faced with this ' logical dilemma,' the investigators are led " to raise the specific issue whether it is an emotional or motivational ***pathology***, or a ***distinctive intellectual ability***, that accounts for the superior scholastic performance." The first alternative—the appeal to ' psychopathology '—they very rightly reject : indeed, in this country few would take such an explanation seriously. The hypothesis which the investigators themselves prefer is that there are two distinct types of cognitive ability, which they call ' intelligence 'and ' creativity ' respectively. " On the perennial question of whether the observed differences are the product of nature or nurture," they say, they would " agree with Guilford on adopting a position midway between the two extremes." In this, once again, they are unquestionably correct : education cannot create creativity ; but it can do much to encourage and develop it.

However, what I want chiefly to question is the notion that there are two separate groups or types each characterized by a distinctive kind of intellectual ability. This method of interpreting the results seems a natural consequence of the factorial theory favoured by most American investigators—namely, that the mind on its intellectual side is composed of a number of ' primary abilities,' and that the factors obtained from the kind of correlation matrix which at first sight suggests a 'general factor,' ought always to be 'rotated' to secure a 'simple structure' of group factors only. Before we accept such a hypothesis we must scrutinize a little more closely the data and the arguments on which it rests. In selecting the so-called ' intelligent ' group the investigators apparently relied on a single ' I.Q.' procured from the "records office of the school." Moreover, the test was applied, not at the time of the inquiry, but when the pupils first entered the school, nor was it the same for every pupil. Hence, the assessments so obtained must have a rather low reliability. And in any case a test of the type employed, as the authors themselves observe in another part of their book, can " rarely account for more than one-quarter of the variance in such crucial factors as school achievement and academic performance." From this they infer that " the common type of intelligence test " fails to "sample all or at least a sufficiently broad range of known cognitive abilities"; it can only represent " a rather narrow band of intellectual activities" ; and, they add, " it thus discourages observation of other types of cognitive functioning."

Nevertheless, this of itself hardly suffices to prove that there is no such thing as a general factor underlying all known cognitive processes. It merely shows that the " common type of intelligence test " provides a very inadequate measure of that factor. But this is a conclusion which has been stressed over and over again in the pages of this journal. Many of the constituent items that make up such tests depend largely on memory, particularly on verbal memory ; they give an undue advantage to the well-taught child from the cultured home ; and they " tend to select children of an analytic or reproductive type rather than those of an intuitive or productive type " (see, for example, this *Journal*, XIV, pp. 135f, and refs.).

Unlike the intelligence tests, those used for testing ' creativity ' were devised and applied specially in the course of the research ; and, instead of just one test, five were used. The authors print a table of intercorrelations for boys and girls, based on 533 cases. Each of the tests of creativity correlates positively with the test of intelligence, two of them to the extent of 0·37 or 0·38. The authors themselves describe these figures as ' fairly low,' and would evidently have us believe that their results fully confirm Guilford's prediction. But the correlations of the various tests of creativity among themselves are not much higher ; moreover, as we have seen, the reliability of the intelligence test must have been rather low, and the children tested were themselves by no means a complete or random sample of the general population. In the table, taken as a whole, all the coefficients are positive ; and, with one doubtful exception among the girls, each of the thirty figures is statistically significant. With both sexes a single general factor would fully account for the correlations observed ; when the effects of the general factor have been eliminated, not one of the residual correlations is numerically larger than 0·10, and none is statistically significant. Indeed, the new tests for ' creativity ' would form very satisfactory additions to any ordinary battery for testing the general factor of intelligence. In the earliest types of intelligence test the child had to invent his own reply ; in most present-day tests alternative answers are suggested, and the child has merely to decide which is correct. This (as has often been pointed out) reduces spontaneity to a minimum ; and an entirely different type of test-problem is needed to redress the balance.

However, when we examine the nature of the tests proposed by Professor Getzels and his colleagues for assessing creativity, and the way in which they are marked, it seems highly probable that they are biased in favour of those particular individuals whose special aptitudes are of a productive rather than a merely reproductive type. And in any case there can be no doubt whatever that these new tests have succeeded in eliciting supplementary activities that are rarely tapped by the usual brands of intelligence test. That this is so can readily be demonstrated by partialling out the influence of those particular processes measured by the 'intelligence' test: we then find that there still remains a set of positive partial correlations between the tests of creativity which are in nearly every case statistically significant.

But, no matter how reliable or how accurate our assessments of the general factor may be, we can certainly agree that this factor by itself could only account for part of " the variance in such crucial performances as scholastic achievement." Motivational factors must be almost equally important ; and special abilities and special disabilities would have considerable weight. In this country, I take it, few educational psychologists would imagine that a single I.Q. based on a ' common type of intelligence test ' would yield all the relevant information required : the real reason why such tests are nowadays so widely used is, not that they are thought to provide adequate or perfect assessments in themselves, but simply because they form the cheapest, most practicable, and most easily standardized means of estimating a child's ' educable capacity.' If, however, as Professor Getzels and his colleagues have apparently found, 'talent-hunters ' in America still trust too exclusively to a single type of test, then we may willingly agree that their criticisms are well founded.

This section of their book ends with a detailed discussion of the theoretical nature of creativity. They briefly review various speculative conceptions put forward by different schools of psychology—the associationist, the Gestalt, the psychoanalytic, etc.; but they entirely overlook the conclusions reached by previous experimentalists in this country. The main results have already been summarized in an earlier issue of this *Journal* (XIX, pp. 180f and refs.). Here, therefore, I need merely note the more important points in which the British studies seem to differ from the American.

To begin with, it has always been a standing principle with British psychologists to provide, whenever possible, some evidence for the reliability and the validity of the tests used. Parallel versions of each new test, for example, are applied to the same group; and, unless the correlation indicates a high degree of self-consistency, the test is usually rejected without further ado. If reliable, the test is then applied either to a group of individuals varying widely in the quality to be assessed and the test-results are correlated with the ratings of a trustworthy observer; or it may be applied to two contrasted groups, selected as possessing a high or a low degree of the quality in question; or, best of all perhaps, a predictive validation may be eventually obtained on the basis of the after-histories of the children tested. All this, of course, implies that we have already decided on a plausible criterion which will furnish a basis for the ratings, the selection, or the later re-assessments—by no means an easy problem in the case of so vague a concept as 'creativity.' Most teachers, for example, tend to limit the term to pupils who show artistic talent or have a gift for imaginative writing: they forget that originality may be equally important in practical or scientific fields of work. Professor Getzels and Professor Jackson, however, apparently expect us to accept their tests as both reliable and valid without any direct or objective evidence.

Nevertheless, here is a collection of ingenious tests, carefully compiled, and the authors may reasonably retort that the critic is free to check the reliability and validity for himself. The British work, by comparison, is still at the preparatory stages. Much of it has been merely exploratory, based chiefly on factorial studies, which, it is assumed, should form the necessary preliminary to the construction of a reliable and valid battery of tests.

The chief conclusions are as follows. First, taking 'creativity' to mean ***useful*** creative activities, most British investigators seem agreed that in all of them general intelligence is an essential, and indeed the most important, constituent. Mere creativity without intelligence, as Galton long ago pointed out, would be all but worthless. Guilford, in a passage quoted by our authors, remarked, that most of the problems set in an intelligence test have 'a unique and an expected answer': "Charles is taller than Henry, John is taller than Charles: who is the shortest?" Here, he says, the examinee's thought-processes are intended to 'converge' on a solution which is already decided, and must be the same for all A test of creativity, say Getzels and Jackson, requires the answers to 'diverge': it measures "the ability to deal ***inventively*** with the verbal or numerical symbols or with object-space relations." But is there any merit in ***mere*** invention—in suggesting that Charles is the shortest, or that twice two are six, just because the answers would be 'divergent'? What I have called 'useful creativity' must involve the ability to deal, not only inventively, but also rationally with the material supplied.

Spearman, indeed, went so far as to maintain that all "creativeness or originality" depends solely on the "eduction of correlates," and is, therefore, merely a manifestation of "general intelligence": "no special creative power exists" (*Abilities of Man*, p. 187). However, nearly every other investigator in this country has found evidence for a distinct group factor—a factor for 'productive' as distinct from 'reproductive imagination.' To that extent the British investigators agree with the American. But they differ from the American in holding that this supplementary factor is not to be regarded as a simple faculty or 'primary ability': it is itself highly complex, and capable of further factorizations into a number of constituent subfactors.

(i) The contributory factor most frequently reported is what Galton termed 'fluency'—i.e., an unusual flow of associated ideas: the creative mind (if I may borrow a phrase from Binet) seems to be "always pullulating with new notions." Further, fluency itself has several different forms—verbal, visual, associative, ideational.

(ii) A second factor, all too frequently confused with the former, is a factor for 'divergent' association (in Stout's narrower sense rather than Guilford's). 'Fluency' might almost be defined in purely quantitative terms—by the number or speed with which ideas are suggested; 'divergence' depends on the qualitative nature of the ideas, and is measured by their variety. In 'serial' or 'redintegrative' association, the whole idea or situation which forms the stimulus tends to reproduce the whole idea or situation which was formerly associated with it; in 'divergent' association only a part of the initial stimulus is operative, and what is recalled may be merely a part of some other situation previously associated with it. Hobbes long ago provided the classical illustration: "in a discourse of our present civil war what could seem more impertinent than to ask (as one did) what was the value of a Roman penny?" (*Leviathan*, I, iii). The operative element in the stimulus was the deliverance of one's master to the enemy; the element recalled were the thirty denarii received by Judas as the reward for the most infamous of betrayals.

(iii) A third component, often overlooked, is a factor for what may be called 'receptivity.' It is this that largely distinguishes the inventiveness of the genius from the spurious originality of the crank, the eccentric, and dreamer. The genius does not merely think of new solutions; he perceives new problems; he sees the familiar in a novel light: he does *not* treat the 'stimulus' solely as a 'point of departure'

(iv) Finally, a factor noted by several investigators is one which they variously describe as 'insight,' 'intuitive synthesis,' or (in James's phrase) 'sagacity.' The true genius does not seize on just *any* distinctive part in the initial situation, nor does he reproduce just *any* distinctive element in the associated idea. He singles out what is most relevant or essential. What may be termed the 'analytic' type of mind—the type that is usually selected by the ordinary intelligence test—commonly reaches the right solution by patient analysis and step-by-step ratiocination; the 'synthetic' type jumps at it by a kind of impressionistic comprehension. Older British writers generally accounted for the process in terms of what Stout and others called 'apperception,' and it was for this reason that 'apperception tests' were introduced into several of the earlier batteries. There is, however, yet a further complication. A number of investigators have reported that several of the factors, usually one or more of the first three, show an appreciable correlation with temperamental factors. The creative child, it is said, is active, alert, and exploratory; and these, it is maintained, are characteristics of an extravertive disposition.

The relative independence of the subsidiary factors I have named seems demonstrated by the fact that they mature (and often decline) at different ages or stages in the development of the child. 'Fluency' appears earliest of all; during the primary school period it frequently diminishes, and revives again in early adolescence. 'Receptivity' may also show itself at a fairly early age. The other two factors have been reported only in studies of older children. Moreover, different factors predominate in differing degrees in different individuals. The creativity of a scientific genius like Newton or Faraday differs widely from that of an artistic genius like Michael Angelo or Beethoven; the fertility of Dickens is different from that of Thackeray or George Eliot; the originality of Browning from that of Tennyson or Blake. For all these reasons, therefore, we must, I think, conclude that the weight of the evidence is strongly against the somewhat simplified interpreation proposed by Professor Getzels and Professor Jackson—namely, that there are just "two basic cognitive or intellective modes," the 'creative' and the 'intelligent,' and similarly, two distinct types of 'gifted students.' Nevertheless, their survey of the subject and their own experimental studies remain full of interesting and suggestive points; and their tests, which are set out in full in an appendix, should be invaluable to future investigators.

The second part of their book describes a shorter study, dealing with what they term ' psychomotor excellence.' The need for some such inquiry, so they explain, was suggested by " the shift in our social and educational goals from values emphasizing ' character ' to those emphasizing ' adjustment '." They contrast " the relatively hedonistic present-time orientation of our culture " and " the traditional work-success ethic " which dominated the " moral climate " of American homes and American schools throughout the nineteenth century. They might perhaps have noted the still more striking contrast between ' present-time orientation ' and the moral outlook of the early Pilgrim Fathers. To demonstrate that much of the change has taken place during our own life time they have selected two specimen periods—(i) between the two wars and (ii) since the last war, and compared the publications listed in the *Education Index.* During the first period they point out the number of references under the heading ' Adjustment ' were only half as numerous as those under the heading ' Character ' or ' Character training ' ; during the post-war period they were six times as numerous, and the references to character training dropped from 369 to only 52.

In their own inquiry they have adopted much the same procedure as before. Two sets of test were constructed and applied. These, it was hoped, would pick out first those pupils who were " high in moral character but not in adjustment " and secondly, those who were " high in adjustment but not in character." The two groups were then compared for intelligence, school attainments, teacher ratings, and home background. In tests of intelligence and of verbal and numerical attainments, the average scores of the ' moral ' group were actually higher than those of the ' adjusted ' group. However, in the intelligence test the difference was too small to be significant. Thus, although the average I Q for both groups was nearly the same, the ' moral ' group had made much greater progress in their school work But some of the children in the ' moral ' group had started school earlier ; and their mothers were sterner disciplinarians and more critical of their children's behaviour. In the occupations and the economic status of the parents there was little difference ; the majority of those in the ' moral ' group lived in ' apartments,' and had changed their residence rather frequently, whereas most of those in the ' adjusted ' group lived in ' houses." There was, too, no discernible difference in the teachers' ratings for ' desirability ; but the number of pupils having ' qualities of leadership,' is said to have been somewhat higher in the ' adjusted ' group.

Once again, however, the groups compared were extremely small—only twenty-seven boys or girls in each, all from the same school Considering that the interviews with each mother lasted " 2 to 3 hours," the information obtained seems rather meagre. Instead of securing factual data relating to the education, occupation, socio-economic class, attendance at church or chapel, or the like, the investigators were primarily interested in what they call " the mother's perceptions of her child." By far the most illuminating portions of the inquiry are the case-studies and the verbatim answers, storyettes, and essays collected from typical individuals in either group. It appears that ' adjusted ' individuals tend, on the whole, to "seek experiences that are immediately gratifying rather than eventually rewarding, to prefer social interaction to personal achievement, and, where there is a conflict of ideals, to sacrifice moral obligations to interpersonal harmony " ; the individuals in the ' moral ' group tend to reverse these trends, and to assert their own autonomy. In short, ' adjustment ' is the ' analogue ' in the sphere of ' psychosocial behaviour ' corresponding to ' intelligence ' in the sphere of ' cognitive behaviour ' ; both are essentially adaptive. ' Character,' on the other hand, is the ' analogue ' of ' creativity ' ; both are distinctive of an original and individualistic type of child.

In view of the limited range of the two sets of inquiries, they must be regarded (to use the author's own description) merely as ' exploratory studies.' Nevertheless, they throw into clear relief two important characteristics which present-day education tends to neglect. Let us, therefore, hope that, in the near future, their investigations will be enlarged and extended, and that similar researches will be planned and carried out by teachers and psychologists in this country.

A REPLY TO C. D. HARDIE'S "RESEARCH NOTES"

BY A. C. AITKEN AND D. N. LAWLEY

In his recent research note, C. D. Hardie attacks the system of inference adopted by most practising statisticians and, in particular, its application in *The Trend of Scottish Intelligence*. It would clearly be impossible in a brief reply to defend this system of inference. Professor Hardie is not the only one to find Fisher's explanations obscure but there are a number of text books by other authors, both British and American, designed to help readers toward a better understanding. To adopt the system of inference propounded by Jeffreys involves, as Hardie remarks, the difficult task of assigning prior probabilities to the alternative hypotheses that intelligence in the population is or is not declining. We find ourselves unable to do this.

In case a contrary impression should have been given, we may say that the mental survey under discussion was very carefully planned, with future statistical tests in mind. Strictly speaking, the 1947 sample was not random. The 36-day sample, for instance, comprised all those Scottish 11-year-old children who were born on the first three days of each month. There was, however, very strong evidence that such a sample would behave as if it were random. The populations about which inferences were made were the populations of 11-year-old children in 1932 and in 1947. The null hypothesis tested was that, for a number of variables, there was no difference between the means of these two populations.

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES*

The Value of the Stern Arithmetic Apparatus in a Backward Class.

BY L. A. IVES

(Abstract of thesis submitted for the degree of M.Ed. in the University of Manchester, 1960.)

AIM OF THE INVESTIGATION.

The aim of this research was to assess the value of the Stern Structural Arithmetic Apparatus to children in a backward class.

THE SUBJECTS.

The subjects used in this experiment were all those, sixteen in number from 7 to 11 years old, who comprised a special class of backward children of low average or borderline defective intelligence in a junior school. At the beginning of the experiment a wide range of standardized attainment tests covering the three R's were administered and all subjects had A.Qs. of less than 85 on all tests

METHODS.

At the outset the subjects were divided into two matched groups, each eight in number. The subjects were matched on three counts, viz. :

(1) Total score gained from three groups of arithmetic tests.
(2) Mental Age gained from the New Stanford-Revision of the Binet-Simon Scale (1937), Form L.
(3) Chronological Age.

There was no significant difference between the groups on any of these counts.

THE ARITHMETIC TESTS.

The Tests Comprising Group One.

It was considered that the tests in this group might be answered correctly by subjects who had learnedb y rote and had not achieved an operational understanding[1] of number.

TEST.

(1) Diagnostic Arithmetic Test, Test 1, Addition (Schonell, 1950).[2]
(2) Diagnositc Arithmetic Test, Test 2 ,Subtraction (Schonell, 1950).
(3) Diagnostic Arithmetic Test, Test 3, Multiplication (Schonell, 1950).
(4) Diagnostic Arithmetic Test, Test 4, Division (Schonell, 1950).
(5) Diagnostic Arithmetic Test, Test 5, Miscellaneous (Schonell, 1950).
(6) Graded Addition Test (vertical form), Writer's own test.
(7) Graded Subtraction Test (vertical form), Writer's own test.

The Tests Comprising Group Two.

The tests in this group tested the chlid's ability to apply his knowledge of the four rules to simple problem situations.

---

* Theses Outlines must be submitted through the Head of the Department in which the research was carried out.

TEST.

8. a. Group Oral One-Step Problem Test—Addition, Writer's own test.
8. b. Group Oral One-Step Problem Test—Subtraction, Writer's own test.
8. c. Group Oral One-Step Problem Test—Multiplication, Writer's own test.
8. d. Group Oral One-Step Problem Test—Division, Writer's own test.
8. e. Group Oral One-Step Problem Test—Miscellaneous, Writer's own test.
11. Group Oral Two-Step Problem Test—Miscellaneous, Writer's own test.

The Tests Comprising Group Three.

The tests in this group tested the child's understanding of mathematical concepts and /or arithmetical processes

TEST.

6. b.[3] Graded Addition Test (horizontal form), Writer's own test.
7. b. Graded Subtraction Test (horizontal form), Writer's own test.
9. a. Individual Test of Cardinal Understanding, Writer's own test.
9. b. Individual Test of Cardinal Understanding, Writer's own test.
9. c. Individual Test of Multiplicative Understanding, Writer's own test.
9. d. Individual Test of Ratio Understanding, Writer's own test.
10. Individual Test of the Understanding which underlies Number Operations., Writer's own test.

THE TEACHING PROGRAMME.

For a period of eight and-a-half months each group was taught by the writer for an hour a day by the ' meaning method.'[4] The teaching programme, which was the same for each group, was based on the programme designed by Stern.[5] The Experimental Group used the Stern apparatus and saw teaching demonstrations with this apparatus. The Control Group used matchsticks and these sticks were used for teaching demonstrations. The Experimental Group used apparatus which *in itself* structures the properties of the number system, whereas the Control Group was not taught to place sticks so as to structure these properties but to use them as aids to computation : thus, the only variable was in the nature of the configurations seen by the respective groups.

RESULTS.

At the end of the experimental period the three groups of tests were administered again. It was found that generally the results increasingly favoured the Experimental Group according to the degree of understanding the child required in order to answer a test, as appears clearly from the following tables.

TABLE 1

THE SIGNIFICANCE OF DIFFERENCES BETWEEN THE TOTAL " (d) " SCORES ON EACH OF THE THREE GROUPS OF TESTS BEFORE AND AFTER THE EXPERIMENTAL PERIOD.

TESTS	In Favour of Experimental Group " t "[6]	Control Group
Group One	1·004	—
Group Two	1·994	—
Group Three	4·003**	—

TABLE 2

The Significance of Differences between the " (d) " Scores on each of the Tests Comprising Group One before and after the Experimental Period.

Tests	In Favour of Experimental Group " t "	Control Group t
1. (timed score)	—	·735
2. (timed score)	·285	—
3. (timed score)	1·783	—
4. (timed score)	·987	—
5. (timed score)	·274	—
6. a.	—	·725
7. a.	2·601*	—

TABLE 3

The Significance of Differences between the " (d) " Scores on each of the Tests Comprising Group Two before and after the Experimental Period.

Tests	In Favour of Experimental Group " t "	Control Group t
8. a.	—	·442
8. b.	2·158*	—
8. c.	1·471	—
8. d.	1·549	—
8. e.	—	·088
11.	1·919	—

TABLE 4

The Significance of Differences between the " (d) " Scores on each of the Tests Comprising Group Three before and after the Experimental Period.

Tests	In Favour of Experimental Group t	Control Group t
6. b.	3·195**	—
7. b.	1·748	—
9. a.	·559	—
9. b.	—	1·038
9. c.	2·653*	—
9. d.	3·245**	—
10.	3·650**	—

GENERAL CONCLUSIONS.

(1) The backward children as taught in this experiment appeared to learn arithmetic satisfactorily when taught by the ' meaning method ' regardless of the apparatus provided, although the children who used the Stern apparatus were superior in both the quantity and quality of learning.

(2) The findings from this research are, in general agreement with those from earlier investigations which have shown structural apparatus to be the most effective aid to learning arithmetic[7], [8]. The fact that these investigations were carried out with normally successful children suggests that the findings from the present experiment might not only be applicable to the learning of backward children, but also to the learning of successful children.

NOTES.

[1] PIAGET (1950) in Chapter V of *The Psychology of Intelligence.* London : Routledge and Kegan Paul.

[2] SCHONELL, F. J. and F. E. (1950). *Diagnostic and Attainment Testing.* Edinburgh : Oliver and Boyd.

[3] Tests 6. b. and 7. b. were made up from the same items as Tests 6. a. and 7. a., but the lines of figures in each item were presented horizontally instead of vertically, viz

Test 6. a. $\begin{array}{r} +27 \\ 25 \\ \hline \end{array}$ Test 6. b. 27—25=

Test 7. a. $\begin{array}{r} +22 \\ 13 \\ \hline \end{array}$ Test 7. b. 22—13=

It was considered that where items were presented horizontally, it was more likely that an understanding of positional notation would be tested.

[4] When the " meaning method " is used instruction is given in an attempt to enable the child to understand the processes he uses rather than work blindly after learning by rote.

[5] STERN, C. (1953). *Children Discover Arithmetic.* London : Harrap.

[6] With 14 degrees of freedom the relevant " t " values are :

5 per cent. level ....2·14 (single asterisk).
2 per cent. level ....2·62
1 per cent level ....2·98 (double asterisk).

[7] MILLER, M. J. (1950). *Investigating the Structural Arithmetic Program of Catherine Stern.* Résumé of unpublished Master's thesis held at Boston Univ. School of Ed., Massachusetts.

[8] KARATZINAS, D., and RENSHAW, T. Testing the effectiveness of the Cuisenaire material. *The Scottish Educ. J.*, 3rd October, 1958.

A Comparison Between Group and Individual Selection Procedures in a Training College.

BY MARJORIE ALLEN

(Abstract of thesis submitted for the degree of M.Ed. in the University of Manchester.)

The purpose of the investigation was to compare the predictive value of different rocedures used in the selection of applicants to a training college.

THE EXPERIMENT.

Different selection procedures were used in two successive years.

(i) *Group selection procedure* (89 subjects). Group techniques were used to show the candidates in a variety of situations, so that as many aspects of personality as possible would be revealed. The applicants met in groups of ten, and took part in leaderless and controlled group discussions on educational topics, mock staff meetings, and short individual interviews. In this procedure considerable weight was given to personality characteristics such a social-maturity, assessed by three assessors who gave individual and agreed ratings to each candidate after each part of the procedure.

The candidates were also given a vocabulary test, a non-verbal intelligence test, and the Allport-Vernon ' Study of Values.'

(ii) *Individual selection procedure* (147 subjects). This was a more traditional form of selection procedure. Each candidate had two individual interviews, and did a verbal and non-verbal intelligence test.

The following year the same procedure was followed (179 subjects). Before their training commenced they did the cognitive tests given the first year.

The main criteria of success were the results of the final teaching practice and the final examinations. Product-moment correlations were calculated between these and the selection test results.

RESULTS.

(1) The final teaching practice results were best predicted by the overall rating of the group selection procedure with a correlation of ·573.

The highest correlation of the individual selection procedure with this criterion was ·478 with the interviews.

(2) The total examination results were best predicted by the average test grading of the individual selection procedure with a correlation of ·569.

The overall rating of the group selection procedure gave a correlation of ·495 with this criterion.

(3) The verbal intelligence test, vocabulary test, and staff meetings were the separate parts of the procedures which gave the best prediction of the compulsory examination results.

(4) The non-verbal intelligence tests were the least useful part of the test batteries.

(5) A selection procedure including long interviews, staff meetings, a verbal intelligence test and a vocabulary test might give a good prediction of both the teaching practice and the examination results, and yet not be an unduly long procedure.

SUBSIDIARY ENQUIRIES

As the students were of different ages (18 to 50 years) and academic backgrounds further analysis of the results was made in terms of :

(i) The entrance qualifications of the students.
(ii) The ages of the students.
(iii) Some discrepancies between the selection procedure and the final examination results.

RESULTS—continued.

(6) Almost one-third of the students admitted to college each year lacked the normal entrance requirements.

In the final teaching practice, there was no difference between the mean scores of these ' exceptional ' students, and students with the normal entrance qualifications.

In the total examination grade the ' exceptional ' students had the lower mean score, and the difference between the means was significant.

(7) The group selection procedure gave a better prediction than the individual selection procedure of the outstandingly good results obtained by the students aged 28–37 years, and of the weakness of some of the younger students.

(8) The vocabulary test did not reveal the weakness in English Language of some of the older students, as well as the verbal intelligence test did.

(9) The young students who failed or had low pass marks in the final examinations seemed lacking in motivation and reliability. The group selection procedure gave a better prediction of this than the individual selection procedure.

(10) Failure of the older students was usually due to a number of causes, lack of adaptability, poor organisation and difficult home circumstances being common to all of them.

(11) Students with low test gradings who did better than expected during the college course were those whose motivation was quite strong, but who lacked confidence. The group selection procedure underestimated these students more than the individual selection procedure.

Social Environment and Success in the Secondary School Selection Examination.

By T. DERRICK

(*Abstract of a thesis submitted for degree of M.Ed. in the University of Manchester,* 1961

AIM.

The aim of the research was to examine certain aspects of family and school environment in relation to the achievement of the daughters of skilled and unskilled manual workers in the secondary school selection examinations in two districts, one in Scotland and the other, the more prosperous, in England.

METHOD.

The city of Dundee and the towns of Leyland and Chorley, Lancashire, were chosen for study. Floud, Halsey, and Martin, in *Social Class and Educational Opportunity* (1957), had shown that there was a tendency for economic differences, when restricted in range, to cease to be the more important factor in educational selection for grammar schools, for boys and for certain other influences, arising from parents' attitudes, to supplant the economic factors in importance. Leyland with Chorley, and Dundee, are two areas with differing economic standards, in countries with the reputation for differences in some attitudes towards education.

From the lists of schools supplied by the authorities, all the junior schools with girls in Leyland and Chorley (eleven) and, in Dundee, a random selection of ten pre-1956 schools from a population of twenty-six such schools, with provision for representation from each group of four school-building standards, were studied. The girls in Scotland are selected for secondary schools one year later than they would be in England. The omission of post-1956 Dundee schools from the sample was done in order to allow sufficient time, it was thought, for the school and home environment to take effect. There were no such schools in Leyland and Chorley.

A guide to the girls' ability and aptitude for grammar or senior secondary school education was obtained from the chosen schools by means of a list of those girls who had, in the opinion of their teachers, a reasonable chance of being selected for such a type of secondary school. The schools also supplied a list of names of a second group of girls, equal in number to the first, and immediately below them in the teachers' rating of this condition. These criteria provided a total of 253 girls out of a sampled school population of 675 in both countries (37·6 per cent), and the parents of these girls were interviewed by the writer in their homes. Those families whose father (or mother, if the father was not living at home) was not a skilled or unskilled manual worker, as defined by the Registrar-General's schedule of occupations in the 1951 Census, were excluded from further study and the resources of the experiment concentrated upon 100 skilled and fifty-seven unskilled manual workers' daughters, who comprised 62 per cent. of those interviewed.

The following variables in the home environment were studied :

(1) Father's Age.
(2) School-leaving age of Father. (3) Of Mother.
(4) School Education of Father. (5) Of Mother.
(6) Further Education of Father. (7) Of Mother.
(8) House type : Old, Modern or Corporation Estate.
(9) House : Rented or Owned.
(10) House : Number of Rooms.
(11) Number of People in Household.
(12) Number of Children in Family.
(13) Occupation of Father (or Mother, if Father deceased).
(14) Basic Wage of Father (or Mother, if Father deceased).
(15) Luxury Possessions : Television Set, Telephone, Refrigerator, Motor Car.
(16) Parent Interviewed : Father, Mother, or Both.
(17) District : Leyland, Chorley, or Dundee.
(18) School : Building standard and condition.
(19) School : Roman Catholic or Protestant.
(20) Intelligence Test Standardised Score.
(21) Result of Selection Examination.
(22) History of any family interest in education.
(23) Length of children's education dependent upon their sex.

Answers to the following questions were recorded :

(24) How often have you thought about secondary education ?
(25) Have you visited your daughter's school during the past year ?
(26) Your choice of secondary school type ?
(27) Reasons for this choice ?
(28) Does this move matter a lot ?
(29) Would you accept the offer of a Grammar (or Senior Secondary) School education for her ?
(For those who expressed no preference in answer to (26)).
(30) A desirable school-leaving age ?
(31) Type of further education desired, if any ?

The criterion was the Education Authority's offer to the parents of the girl concerned of a grammar school or senior secondary school education (defined as a 'Successful' girl), on the one hand, or a modern or junior secondary school education (defined as an 'Unsuccessful' girl), on the other. In Dundee, a further class was distinguished—a small number of girls who were offered senior secondary school places on probation, because of the Authority's uncertainty about their aptitude or their ability or both.

The standardised intelligence test score was the result obtained by each girl in her selection examination. The tests used in both districts were Moray House, standardised on 'outside' populations.

Certain items from the schedule of questions, concerning attitudes, were calculated on an arbitrary points scale and categorised from 1 (Low) to 4 (High); similarly with Home Standard items.

Comparisons between the constituents of the tables of results were calculated for significance by Chi-Square, or by the 't' test where appropriate.

RESULTS.

The following significant differences were found. The first listed has the superior standard in each case :

Housing

Lancashire/Dundee, $P. < \cdot 001$.

School Buildings

Dundee/Lancashire, P. < ·001.
Dundee Protestants/Lancashire Protestants, P. < ·001.
Dundee Roman Catholics/Lancashire Roman Catholics, P< ·001.

Home Environment

Lancashire/Dundee, P. < ·01.
Lancashire Skilled/Unskilled, P. < .001.
Dundee Skilled/Unskilled, P. < ·01.

Preference for Grammar School

Lancashire Skilled/Unskilled, P. < ·05.

School-Leaving Age Preferred

Lancashire parents preferring Grammar Schools/Modern Schools, P. < ·001.

Size of Family

Dundee/Lancashire, P. < ·01.

In the tables which included the criterion, the following significant differences were found. The first listed has the superior success rates :

Intelligence Test Scores

Lancashire.

Skilled/Unskilled, P. < ·01.

Differences between the parents of Successful and Unsuccessful girls were found in the following :

Preference Leaving Age of 18 *Years*

Dundee, P. < 01.

A History of Family Interest in Education

Lancashire, P. < ·01.
Lancashire Skilled, P. < ·01.
Lancashire Protestant, P. < ·01.

The Occupation of Older Houses

Lancashire, P. < ·05.
Lancashire Skilled, P. < ·05.
Lancashire Protestant, P. < ·05.

The Ownership of Houses

As for Older Houses.

The following district difference is an initial requisite for the appreciation of the significant differences stated below :

Distribution of Awards of Grammar School Places

	Chi-square	d.f.	Sig.
Lancashire/Dundee	1·02	1	N.S.

Successful Girls

Awards of Grammar School Places in relation to :

The Further Education of the Parents

Lancashire/Dundee, P. < ·02.

The Occupation of Older Houses

Lancashire/Dundee, P. < ·05.

A Study of Personal Attributes and Roles Favoured by Boys and Girls Aged Eleven to Thirteen Years.

BY D. H. BRAY

(*Abstract of a thesis submitted for the degree of Ph.D., University of London*, 1961.)

AIM AND METHODS.

With the aim of ascertaining reasons for the favouring of, or identifying with, people, a study was made of attitudes to characters in stories on the part of English school-children around the age of entry to secondary education: a questionnaire was given on a play that had been seen, stories were written on set topics and, following incomplete stories presented in dramatic form, completion and other responses were asked. In addition, questionnaire information, not limited to narrative reference, was sought concerning desired attributes and roles and sources of acquaintance with favoured people, e.g., through television.

INCOMPLETE STORY RESPONSES.

As the main enquiry, six incomplete stories forming a sequence were dramatically presented weekly at the end of each of which structured completion and other responses were asked in order to determine the favouring of characters who were distinguished by postulated qualities and roles. Each completion questionnaire included three standard questions requiring choices from among the characters: (1) participation in an outcome postulated as favourable; (2) the conceiving of one of the characters as relating afterwards the events of the story; (3) the preferred part for acting by the subjects. The subjects, Form II pupils with a mean age of 12 years 8 months, were sixty-six boys and seventy-eight girls (of whom forty-one and thirty-eight, respectively, were present throughout the series) in two co-educational secondary modern London schools.

As influences fostering the favouring of characters, subjects' similarity to characters in age or in national group received no support from the results but there was partial support for similarity in sex. The evidence showed the importance of needs imputed to the characters. Thus, following the presentation of his needs for parental acceptance and for achieving a desired vocational role, a hitherto least favoured character became so favoured that out of five characters he alone was significantly chosen on any of the three standard questions already described: he was significantly selected by boys and by girls at the ·01 level on all three of these questions. That another character enacting a leadership role was highly favoured in the story series by frequent preference at the ·01 level, may similarly be inferred to relate to the satisfying of the needs for acceptance and responsible contributing to a group.

A significant tendency to support familiar foreign characters against British newcomers may be attributed to the potency of stories in building identifications with characters. The occurrence of identification might also be inferred from significantly greater constancy of character choice by a group of boys instructed to role-take (i.e., to imagine themselves each as enacting one of the character roles) as compared with boys not so instructed.

Validating support for the Incomplete Story procedures included: opinions of judges in matching stories with hypotheses and expected responses, accordance of subjects' opinions with postulated roles and attributes, significantly successful outcomes desired by the subjects for chosen characters with the implication that the choosing involved favouring and, in relationship to degree of constancy of character choice, both preference for people on the Cotswold Personality Assessment: P.A.1.* and role-taking as a controlled variation in procedure.

---

* FLEMING, C. M. Glasgow: R Gibson and Sons, 1959.

OTHER ENQUIRIES.

The favouring of characters in a narrative context was also studied in three further enquiries: Responses to a Play where, following the presentation of a dramatization of Hans Anderson's *The Snow Queen*, thirty boys and sixty-nine girls who had voluntarily attended answered questions aimed to determine their favouring of characters; Preferred Topic Story Writing where the writing of stories by 218 boys and 216 girls on their choice from set topics was analysed; Secondary Selection Essays, where two samples of, respectively, 1,252 and 174 boys' and 1,163 and 180 girls' essays were studied from the results of an 'Eleven Plus' examination in which was set the compulstory topic of writing part of a biography of an admired adult. Lastly, in a Self and Others Questionnaire, eighty-five boys and ninety-nine girls were asked about whom they would most wish to resemble, how they thought their peers imagined themselves and concerning communication sources for people they especially favoured, e.g., school, television, cinema, etc.

The findings from these enquiries again showed the importance of needs: the favouring of characters was associated with the needs of acceptance, success and its relationship to status and ability, responsible contributing to a group, the giving and receiving of help and friendliness.

In all four of these enquiries there were, too, significant tendencies for the favouring of characters to be associated with likeness in sex to the subjects but the tendencies were more pronounced for the boys than the girls who in their choices of remote, though not of local, admired adults as topics for the Secondary Selection Essays chose males as frequently as females. In this same enquiry there was a further tendency to sex differentiation of choice as in girls' relatively greater writing about admired characters classified as humanitarian and religious together with greater stress on the giving and receiving of affection compared with the boys who chose relatively greater numbers of characters associated with war and politics, exploration and sport.

By both boys and girls, admired adults tended to be chosen from the remote environment. In the Self and Others Questionnaire, girls tended to choose other local children significantly more than the boys who had a greater preference for admired adults. However, the considerable extent to which both boys and girls chose adults was in accordance with the general failure to find similarity in age to the subject as a favouring influence. In this same enquiry television as compared with eight other communication media was significantly stressed (·01 level) as a means of acquainting the boys and girls with people in whom they were especially interested.

The Preferred Topic Story Writing results included a finding relevant to the concept of identification: the degree to which the first person was used by the main character in relating the story varied from much usage in attractive topics (e.g., 'allowed to do anything one wants') to no usage in the extremes of rejection involved in being 'badly treated.'

DISCUSSION.

In answering a question raised by the Himmelweit Television Enquiry as to the determinants of identification,* the results have tended to stress the importance of the subjective and common characteristics, needs, as contrasted with characteristics that are more 'objective' and that pertain to differences in people, as age and nationality, though sex of characters was shown to be of some importance.

From evidence supporting its validation and from productivity of significant results, the story completion method is indicated as holding promise for further research.

ACKNOWLEDGMENTS.—The helpful guidance in this investigation by Dr. C. M. Fleming is gratefully acknowledged. Thanks are also due to two further members of the University of London Institute of Education staff: to Mrs. D. M. Edwards Penfold for statistical advice and, in conjunction with the Secondary Selection Essays

---

* HIMMELWEIT, H. T., *et al.*, (1958). *Television and the Child.* London: Oxford University Press, p. 404.

to Dr. W. H. King. Indebtedness is acknowledged to the Imperial Relations Trust and the New Zealand Government for the award of a fellowship and for leave and to the Central Research Fund of the University of London for a grant in aid of this research.

SUMMARY OF THE CONCLUSIONS.

1.—The two districts differed significantly on certain items. Lancashire had superior living standards and owned far more homes. Dundee had larger families and superior school buildings.

2.—There was no such difference between them in Parents' Attitudes. Both showed strong preferences for grammar or senior secondary schools, a school-leaving age of 18 years, and a university education for their daughters.

3.—Skilled parents had better living standards than did Unskilled, and sometimes more favourable attitudes towards the education of their daughters.

4.—Religious groups (Protestants and Roman Catholics) did not differ on these scores except for the following:

(*a*) Roman Catholic children were in school buildings inferior to those of Protestant children, and

(*b*) there was a tendency for Roman Catholic families to be bigger than Protestant ones.

5.—Enlisting the support of education authorities in this research was difficult in England: enlisting the co-operation of the parents was not.

6.—In a comparison of environmental factors with the criterion (Pass/Fail for admission to a grammar or senior secondary school) no single item was significantly related to it in both districts. In Lancashire there were significant relationships with (*a*) a history of Family interest in Education, (*b*) Ownership of Houses, and in Dundee with a desire for Further Education. Other items which had the appearance of being related, in both districts, were (*a*) Grammar School Preference, (*b*) No Sex Discrimination in the length of School Life, (*c*) Allocation to Secondary Schools very important, and (*d*) Father aged 50 years or more.

(7) The difference between Skilled and Unskilled Parents in the proportion of awards to Grammar Schools in Lancashire is insignificant.

8.—Floud found that the children's success rate in Middlesborough suffered from economic hardship in spite of parents' educational aspirations. Dundee children did not show this definite trend. In Lancashire, wages did not correlate with the criterion. "Owned Houses" did. It may be that ownership of a house is not mainly a sign of economic standing—as Floud interpreted it—but depends on an attitude of mind, amongst working class parents.

BOOK REVIEWS

BARRINGTON KAYE (1962). *Bringing Up Children in Ghana.: An Impressionistic Survey.* London: Allen and Unwin, pp. 244, 40s.

There are several valuable studies of child rearing in African communities, undertaken by European anthropologists and psychologists. The material offered by Dr. Kaye is the first collected by Africans themselves. Students at the Institute of Education were required, as part of their course, to write a report on child training in their home town or village. For this purpose, they were provided with an outline of topics to be covered, and this forms the framework for the presentation. After little more than half a dozen pages on the general social and demographic background, the reader is plunged straight into an account covering all the major phases of growth from conception to adolescence. The author's method is to string together long series of quotations from the reports, linking them with brief comments and referring to available data from anthropological studies. The resulting impression, though very rich and colourful, is perhaps apt to be somewhat bewildering for those without first-hand experience of life in Ghana.

Devotees of Dr. Spock, and more generally those who believe that we possess firm knowledge about the relationship between child-rearing and adult personality, are liable to be upset by this book. After a nursing period of almost complete permissiveness, during which the child is extremely close to the mother, there often follows harsh and abrupt weaning, with the mother applying bitter herbs to the breasts to discourage the child. Methods of controlling behaviour are frequently drastic by our standards: apart from beatings, there are threats of bogies, wild animals or ghosts; for some offences, like early sex play, pepper is rubbed on the genitals (though there is considerable sexual freedom from adolescence onwards); children are told not to touch their nose in the course of a meal, lest a parent die as a result. What sort of adults do these children become? We are not told anything about that. In the reviewer's opinion they are, on the whole, easy-going and " well-adjusted," though not without their anxieties; but then, who is?

Kaye is careful to disarm criticism at the outset by disclaiming scientific rigour; nevertheless, a few questions ought to be raised. The first concerns the legitimacy of referring to child training in " Ghana," which until comparatively recently was little more than a geographical expression, and certainly remains very heterogeneous in terms of tribal cultures and such factors as literacy and urbanization. In fact, even these edited reports indicate a wide range of variation; moreover, there is no means of telling how far they reflect real differences, and how far merely the bias of untrained investigators. No attempt is made to pull the material together, indicating common and divergent trends.

In view of the unsystematic nature of the data, the author's hope that they may help to generate " hypotheses " is somewhat questionable. The book does, however, provide many qualitative insights into approaches to child training which, in many respects, differ radically from our own. Educationists, psychologists and others concerned with child development will find this rewarding. All readers can enjoy the numerous photographs taken by the author; these are not only technically superb, but have the rarer virtue of being really pertinent to the text.

GUSTAV JAHODA.

EVANS, K. M. (1961). *Sociometry and Education.* London: Routledge and Kegan Paul, pp. 149, 18s.

The recognition of the influence exerted by the group on the individual and interest in the structure and development of groups has been accompanied by the development of techniques for gathering and recording reliable data on these topics. They are grouped together under the generic heading of Sociometry, although this title is often used in the more restricted sense to apply to techniques which have derived from the work of J. L. Moreno and involve the use and interpretation of sociometric tests.

Because such tests are easy to apply and require no advanced knowledge of statistics they have come to be a favourite tool of the research worker in the social, sciences and have been widely used in education both in this country and in America. For this reason an account of their construction and use and a critical appraisal of their findings is very welcome. The first part of Dr. Evans' book is devoted to these topics and teachers will find a great deal to interest them in the results of sociometric studies which have investigated aspects of popularity and leadership in a school setting. Two points might be made about this section. Firstly, the teacher who wishes to use a sociometric test would have been helped by the inclusion of an actual example. The form of items is mentioned, cautions are uttered concerning their relevance to what is being measured, and some examples of sociograms are given, but it is too easy to move into research in this field, without considering what objective characteristics of behaviour are being examined and to suffer from ' diagram dazzle ' at the end. Secondly, whilst teachers will be quick to recognise the value of applying the information revealed about interpersonal relationships, they may find their suspicions confirmed that the social sciences tell us little that we do not already know. To overcome this, long-term studies would be required which would demonstrate that the manipulations made as a result of applying sociometric techniques have some lasting effect in terms of learning and adjustment. Dr. Evans cannot be blamed for the absence of such studies but more might have been said about the failure of many investigators to relate sociometric measures to fundamental variables in a systematic way.

The later chapters of the book are occupied with a more general discussion of problems of communication and social perception which have been investigated by means of rating scales, personality inventories and questionnaires. Although these have been conducted in a variety of settings, the emphasis is upon those findings which have most relevance to the classroom. Here again, studies reveal some interesting information about the way in which teachers regard pupils and *vice versa*, yet we know comparatively little about the variables which determine these impressions and whether social perception follows the same laws as the perception of non-social objects.

It is the absence of central hypotheses of the kind mentioned that makes a record of the empirical findings in this field somewhat tedious reading. There is no doubt, however, that teachers and students will find much in this book which will enlarge their understanding of classroom relationships. P. K. POPPLETON.

GELDARD, F. A. (1962). *Fundamentals of Psychology*. New York : John Wiley, pp. 437, 57s.

The series of introductory text books by Boring, Langfeld and Weld, have, since the first one appeared in 1935, been justifiably popular with psychology students. Their aims have been similar to those of the standard texts in the physical sciences at this level in that they seek to present to students the generally agreed facts in the whole field of the science ; in this series notional beliefs and speculative theory have had no place. Changes in emphasis and increments of knowledge between revisions have been such that each new B.L.W. was so substantially different from its predecessors that it warranted a new title. Thus, the *Factual Textbook* of 1935 became an *Introduction to Psychology* in 1939 and *Foundations of Psychology* in 1948. The book under review is a direct descendant of the same distinguished line but even more than the previous texts it is a complete revision and bears little superficial resemblance to them. Its aims are, however, the same, and its plan was sketched some ten years ago by a steering committee which included the three original editors. This time, however, the book has been written completely by the one author, Professor Geldard. It has gained in that the central plan is more coherent and the writing more even.

The arrangement is cyclic. After a short introductory chapter on the nature and scope of the science of psychology, there follow sections which deal at a basic elementary level with the three major aspects of motivation, learning and perception. The first of these is presented clearly in terms of need and drive, it includes discussion of innate behaviour patterns, such as instincts, reflexes and tropisms, and of the emotions as motivators of behaviour. Learning, in this early part of the book, is dealt with only in terms of classical and operant conditioning. Elaborations of behaviour are seen as the results of stimulus generalisations, secondary reinforcement. higher order conditioning and transfer of learning. Finally, in this first part of the book perception is described as the essential procedure of gathering information through the sense organs.

When these basic elementary concepts have been presented, the author feels free to return to them again and to elaborate and develop them. Now the social motives are introduced, learning is taken well beyond conditioning and is seen as, for example, the development of skills, remembering and forgetting, problem solving and thinking. The idea of perception is extended to include perceptual organisation, attention and set, constancy phenomena, perception of space and time and so on. This leads quite naturally to such topics as personality, abilities and group dynamics which are all dealt with very adequately for an introductory text, emphasising general principles and accepted facts and complete enough to challenge the good student.

This is a worthy successor to the *Foundations*. It is much less verbose than many American elementary texts being well written and well presented. The diagrams are clear and relevant, but it contains a lot of photographs of long dead psychologists which seem to add little to the value but a lot to the price of the book. The question must be asked as to whether it would do as the main text for teachers in training. Probably it would not for its setting is general, its approach to learning via conditioning would seem remote from the classroom and it pays scant attention to British work. Nevertheless, if it were read selectively and was adequately supplemented by discussion and further reading so as to ensure transfer of the general principles to classroom situations, it could form a useful basis of study, specially with more able students. It should certainly be in the library of colleges and departments.

L. B. BIRCH.

GILLETT, A. N., and SADLER, J. E. (1962). *Training for Teaching*. London: George Allen and Unwin, pp. 272, 21s.

Almost a quarter of a century ago, P. B. Ballard, in this *Journal*, reviewing a book on the general principles of teaching, made the point that there were two schools of thought in education, the philosophic and the scientific: " One of which concerns itself with discussing first principles and the other with discovering facts and formulating laws." He went on to add that without the scientific school progress in educational practice would be impossible and that, " educational theory becomes mere bombination in the void." He further asserted that presented with a book claiming to cover both schools one may legitimately raise a protest. This is that kind of book and the authors appear to be quite unaware that the kind of distinction Ballard made is one which can and ought to be made.

When the authors claim, as they do in the introduction, that this book is intended to be a guide for the student in training, they must simply have had in mind that a book's parts, chapters, index, bibliography, subjects for discussion and page numbers technically qualify it as a guide. So they do, but they can hardly be expected to guide the student to an independent appraisal of the nature of educational theory and a critical assessment of educational practice. Something more is needed than a potted history of education, an introduction to educational philosophy which omits Rousseau and Dewey, and brief accounts of learning theory, human abilities and Piaget's work.

Throughout the book the authors make judgments of the kind " teachers should," " children ought," " society must " and so on. They also make the udgment that teaching is a service and that teachers need all the help they can get from educational theory (*singular*). They hardly mention the educational research of the last fifty years and give no space to the application of the methods of the social sciences to education.

This book is no guide. It follows in the hallowed tradition of so many books in education, some great, some not so great, which have attempted to provide students in training with harness and blinkers which will fit them to pull their weight unflinchingly through the years of teaching service which lie ahead of them.

The authors write with sincerity, with moral fervour but without real awareness that what is required to-day in education is not more judgments but a more rigorous examination of those with which the past has saddled us. P. H. TAYLOR.

GRUBER, F. C. (Edit.) (1961). *Anthropology and Education.* The Martin G. Brumbaugh Lectures in Education, Fifth Series. University of Pennsylvania Press and Oxford University Press, pp. 123, 22s.

In a broadly based lecture series on education, there is an obvious case for asking anthropologists to contribute. As a realisation of what is in itself a good idea, the contributions in this small volume are, though serious-minded, disappointing. It is perhaps inevitable that in leaving the confines of their own discipline, these lecturers should have taken the opportunity of trying out some of their pet ideas in social theory and philosophy. They raise big questions, of the kind which all academics and other intellectuals of to-day find worrying, and in a way which must have been stimulating to their audiences when the lectures were delivered. But read at a distance, in cold print, their answers seem as platitudinous as speeches at a prize-giving.

Professor Wallace does, indeed, arouse our interest when he writes : " To-day, our society here in the United States is conservative. We went through our major revitalisation movement nearly two-hundred years ago." He goes on to "suggest what, in a conservative society intending to survive in a revolutionary world by refusing to freeze into the reactionary posture, the value hierarchy of the schools should be. The cultivation of intellect should come first, technic second, and morality last." This may well be a good comment on widely discussed deficiencies in American education, such as any self-respecting intellectual (but not necessarily an anthropologist) might feel inclined to make.

Professor Hymes is an anthropologist with a special interest in linguistics, and in his lecture he advocates an evolutionary approach to the study of language, or rather a return to the evolutionary approach, now that its early errors have been discarded. He refers to significant research into the attitude towards their language of various ethnic minority groups in North and South America. Such attitudes may, of course, be highly relevant for educationalists in relation to such peoples. But the case for the general relevance of this kind of study for education is not fully made out.

Professor Goodenough's theme is the kind of educational problem that arises in community development programmes " where teacher and pupil represent different societies and cultures and where change in the pupil's way of life is the educational goal." How is such change to be brought about without loss of self-esteem, national and individual ? He cites the case of the Makanai people of New Britain Island, among whom he worked, and their enthusiasm for corrugated iron roofing as a " symbol of the European way of life," although, being uninsulated, it made their houses impossibly hot. Europeans who tried to dissuade them were suspected of being opposed to their wider aspirations. This, is, indeed, a delicate problem, to be found throughout the tropical world, and made more difficult by the rational and economic arguments in favour of metal roofing and by the irrational element of aesthetic sentiment in the preference of Europeans for the use of thatch.

Dr. Dorothy Lee reads a lesson to the teachers on their over-concern with competitive achievement. Her contact with the more generous motivations of non-Western cultures has led her increasingly to question the utilitarian assumptions of the West. She admires the destitute Arabs of the Empty Quarter who pressed food on uninvited guests, though they had themselves had nothing to eat or drink for a day, and would have nothing for another twenty-four hours. " This," she comments, " is excessive giving, beyond the limits of common sense or rationality," yet she deplores the absence " in our own culture " of any " official recognition of exuberance." Awareness of such un-American values is perhaps something which anthropologists can valuably bring about in American pedagogues. But it seems a little surprising that a lecture on the value of self-sacrifice should be needed from this quarter. CHARLES MADGE.

HALSEY, A. M., FLOUD, J., and ANDERSON, C. A. (Edit.) (1962). ***Education, Economy and Society.*** New York and London: The Macmillan Co., pp. 625, 56s.

Some little time ago Diogenes pronounced that, " The foundation of every State is the education of its youth." To say that this book supports Diogenes would be truthful but critical by implication, and the latter would be ill-deserved. There are many books on education which merely place Diogenes and other worthies in a modern setting ; what the former said in his tub the present pundits echo from their ivory tower. Such books have their value in reminding us of fundamental aims but they add nothing to our knowledge. Books on educational psychology and educational sociology, on the other hand, are often at the growing point of education. This is such a book. It is a judicious collection of articles by many writers ; reprints of articles written by the editors themselves constitute almost a sixth of the text, though one of these chapters is written jointly with another writer.

A principal theme of the first section is that of Diogenes, but here it is well supported by facts and illustration. Entitled " The Consequencies of Economic Change," it deals with the interaction between education and society in a quickly changing milieu. In the words of the Introduction, " No longer is it a question of handing on an unchanging body of knowledge and belief ; on the contrary, education in modern societies is more concerned with changing knowledge than with conserving it." A powerful chapter by Drucker illustrates vividly the explosive nature of the educational revolution in U.S.A. and Soviet Russia ; it should be compulsory reading for every member of the Cabinet. " An abundant and increasing supply of highly educated people has become the absolute pre-requisite of social and economic development in our world. It is rapidly becoming a condition of national survival." Later, Ginsburg contrasts the American willing acceptance of economic change with the resistance to such a change which hampers development in many sectors of the British economy. Another interesting chapter, especially for educational sociologists, is that by Turner on " Modes of Social Ascent through Education" Unfortunately, Marshall, writing on " Social Selection in the Welfare State," appears to accuse psychologists of separating children into three types of ability, viz. ' g ' or general, ' F ' or technical and ' K ' or spatial. This is absurd. Marshall should know that the body which boldly classified children into three types was the Norwood Committee, which contained not a single psychologist, and whose report was roundly condemned by psychologists mainly because of this very division. An unusual article by Anderson endeavours to show that education is not such a dominant factor in social mobility as many sociologists have tried to indicate ; it might have been illuminating, however, to have known the ages of the sons when their social class was assessed—there might yet have been time for them to have moved further upwards and to have surpassed their fathers.

In an important section on " Social Factors in Educational Achievement," Nisbet's article stands out as an able and well written piece of work, with sustained logical reasoning. On only one point would the reviewer question his thesis. In attempting to determine whether a negative correlation exists between verbal ability and family size, independently of any association with intelligence, Nisbet appears to take no account of social class : clearly, however, both size of family and extent of language development are linked with social class, and the consistent drop he gets for test after test as the family gets larger could be due to social class alone.

This is not to deny that the negative correlation probably exists even after social class has been taken into account. The following chapter, by Bernstein, is a valuable complement to that of Nisbet.

Only brief mention can be made of other points of interest to psychologists. Kahl provides more evidence of the influence of social class on attainment (4th to 9th grade), and that in the ' common man ' group it was parental pressure and aspiration which produced the difference in the student's level of aspiration. Coleman examines the influence of teaching methods involving student co-operation on attitudes and attainment, Mead writes of the " nightmare reversal " in modern life, where " age and experience become not orientating factors, but disorientating ones, so that a teacher of 20 years experience may face her class less confidently than the teacher with only 2 years." Three chapters on the universities are of particular interest. Ashby shows clearly how the resistance in Britain to the inclusion of science and technology in the traditional curriculum in the nineteenth century caused Britain to lag behind in industrial competition ; the delaying hand now is that of the Government. Rieseman writes a long and informative article on American universities, while Caplow and McGee give a vivid account of the rat race amongst university staffs over there—an article which will give intending British academic emigrants much food for thought.

This is an important book which should be on the reference shelf in the library of every educational psychologist and educational sociologist. R. R. DALE.

HARING, N. G., and PHILLIPS, E. L. (1962). *Educating Emotionally Disturbed Children*. McGraw-Hill Book Co., Inc., pp. ix+316, 50s. 6d.

This book describes an experiment in the education of emotionally disturbed children conducted in the elementary schools of Arlington, Virginia, U.S.A It is interesting as one of the few attempts to carry out a controlled experiment in the field of maladjustment and to evaluate the results. Nine special classes for emotionally disturbed children were set up in normal primary schools. In three of these the methods used were those advocated by the authors, in three others the methods were those normally used in American schools and in three the methods were what is commonly known as ' permissive.' At the end of the first year tests of academic achievement and ratings of behaviour were carried out, and where possible, similar evaluations were made during the second and third years of the experiment.

In every case the authors claim superiority for their own methods. Their attitude to the problem of maladjustment is, however, essentially educational. The word ' psychiatry ' does not appear in the Index and on the one occasion when a parent, in group discussion, asked whether her child might not benfit from psychotherapy, she was told that therapy would probably not only be unhelpful but might be positively harmful. The reason for this again is that therapy is essentially permissive, whereas the essence of the authors' methods is that they call ' structure.' The theory behind this is apparently a Behaviourist one, as there is a good deal of talk about ' extinction ' and ' interference.' The methods used were largely those advocated by Strauss and Lehtinen for the education of brain-damaged children and their apparent success may have been because the maladjusted children in these special classes were also, like the brain-damaged, hyperactive, and presented relatively simple problems of discipline and learning. The readjustment of the children was to be brought about by a carefully planned and firmly structured programme in which there was no loophole for disobedience or for badly-done work. The parents of the children were required to carry out the same strict regime at home and frequent conferences were held in which each child's behaviour and school work were discussed and detailed advice was given by the psychologists in charge, to both parents and teachers. This account of the work done with parents is one of the most interesting parts of the book. The book is also of value in that it underlines the damaging effects on a child's growth and happiness of mental muddle and lack of direction. It may be that the methods described here would be extremely useful with a certain type of maladjusted child. E. M. BARTLETT.

JACKSON, BRIAN and MARSDEN, DENNIS (1962). *Education and the Working Class*. Routledge and Kegan Paul, pp 268., 28s.

This is without doubt the most exciting essay in educational sociology that has yet appeared in this country. The authors were pupils at a North Country grammar school who conceived the idea of visiting a sample of former boy pupils of working class origin who left school at 18 in the years 1949 to 1952, and who had reached an age between 25 and 30 ; and a similar sample of girl pupils from the sister school, drawn from a somewhat longer period. They also visited the parents of the children in their sample, which totalled 88.

Broadly, their aim was to find out how successfully the children now adults, had adjusted to the changes in social and occupational orientation implied in their education. They seem themselves to have been motivated by a strong dissatisfaction with the predominantly middle-class values of the school, and their attitude cannot in any sense be regarded as detached or ' objective.' Yet, their involvement in the subject-matter gives a bite to their research and to its conclusions which it might otherwise have lacked. This is sociology at its most readable.

Of course, the Marburton schools may not have been typical, and Marburton may have had class norms of a peculiar kind. Only comparative studies could tell us. Another possible criticism is that the years immediately after the war were educationally disorganised. Unfortunately, whatever time or place had been chosen would have had disadvantages of some kind. One must take the evidence presented for what it is and not credit it with a ' scientific ' validity or generality which no such limited survey could possibly confer.

The story of the eighty-eight children is traced, with sensitive skill, from their primary school days down to their fully adult status. Twenty girls and one boy went to training college ; sixteen girls and thirty-eight boys went to university. Almost all are now ' professional class people,' over half of them teachers. " To-day, most of these eighty-eight children have developed into stable, often rigidly orthodox citizens, who wish to preserve a hierarchical society and all its institutions as they now stand." They were, in any case, predominantly drawn from good ' mixed ' social environments, from that part of the working class which is in its values and standard of living nearest to the middle class. An appendix on a small subsidiary sample of early leavers is a useful reminder of some of the ways in which other children, with more ' unmixed ' working-class background, are, so to speak, reclaimed by their first environment. One would welcome a more extended investigation of this group.

This is a book too packed with interest to summarize and the most useful function for a short review like this is to urge that it be read. CHARLES MADGE.

JOHNSON, J. C. (1962). *Educating Hearing-Impaired Children in Ordinary Schools*. University of Manchester Press.

For years research into the deaf has concentrated on their discovery, assessment, classification and placement. The measurement of their hearing and the perfection of apparatus for amplifying sound so that they may hear better have occupied the attention of more and more scientific workers. Those interested in the *education* of the deaf might well turn with relief to a book with the title of this one, hoping for a treatment of problems of education such as what and how to teach, what aims to hold and what place in society to achieve : alas, no. There is little or nothing here about education in this sense. But it is a very interesting research for all that and ought to give those concerned with teaching the deaf much to think about, whether they work in special or in ordinary schools. Moreover, it ought to be studied by education lecturers in training colleges and university departments of education whose students will find children with some degree of deafness in their ordinary classes. The failure of the present ordinary training course to equip teachers to detect defects in their pupils has rarely been brought out so well.

In days gone by there was only one place for the deaf child—in a deaf school. After the 1944 Act, following the advice of a Board of Education Committee of

1938, whose report Mr. Johnson refers to several times in an approving way, a few boarding schools for the partially deaf were set up, and soon after, ' units ' for partially deaf children were opened in ordinary primary schools by several L.E.As. Mr. Johnson has gone into new territory by investigating the hearing, speech, intelligence and attainment of children with medium to severe losses (over 30 db) who are not receiving any special education while they remain in ordinary schools in one English county. A pretty picture of past neglect emerges in spite of the School Health Service, the E.N.T. clinics and the Medresco transistor hearing aid. It is to be hoped that the project, of which Mr. Johnson's research was part, referred to in the preface by Sir Alexander Ewing, causes at least one L.E.A. to improve its facilities.

The details of the picture are carefully set out in Chapters I to VI, and the methodology appears sound. (But why will people still use word-reading tests instead of comprehension tests in assessing retardation ?) Chapter VII gathers together some simple advice to teachers and parents and school authorities which could lead to the avoidance of neglect and the prevention of some backwardness. Next, there comes a quite unrelated Chapter VIII on " Units " which seems to have no research structure behind it and, therefore, is no more than Mr. Johnson's views on a facility which did not exist in the county in which he worked. The views are those of someone who has little understanding of the difficulties of the heads of the host schools.

Tucked in Chapter IX, a kind of afterthought, is a study of 76 of the 120 deaf children who, before he began his work in the county, had been sent to special schools outside the county, and here the sparks fly. The policy of separating the deaf and the partially deaf advocated in 1938 is shown to be still incompletely carried out. Partially deaf children in special schools for the deaf and partially deaf are shown to have lower reading and arithmetic ages than children with the same degree of hearing loss and intelligence who are in ordinary schools even under the unsatisfactory conditions set out in Chapters III to VI. If this were the whole story (and many people could guess that it is not), then why not close up the special schools and put the children back in the ordinary schools, not even in ' units ' ? Mr. Johnson clearly doesn't want that, but he doesn't pursue the matter, asking only for further research. In his last chapter, he goes back to his investigation in the ordinary schools where he is on solid ground, with the advice : " The child with a significant hearing-impairment in an ordinary school, whether there by accident or design, cannot manage without some form of special help."

Definitely a book to be studied.

JORDAN, THOMAS E. (1961). *The Mentally Retarded.* Columbus, Ohio : Charles E. Merrill Books, pp. x+355, 50s.

This book surveys a wide range of theoretical issues and practical problems in the field of mental retardation. The author writes from the viewpoint of a psychologist and professor of education. In addition to the expected chapters on the nature and sources of mental retardation, the characteristics of the mentally retarded and on psychodiagnostics, there are chapters on therapeutic considerations, education and rehabilitation. One chapter gathers together information about the family problems—the effect of the retarded individual on the family and the need for parent education and counselling. There is a final chapter on cerebral palsy and there is a useful survey of medical factors in retardation.

The basis of the book is an extensive survey of the literature of mental retardation and allied fields. While references to American publications predominate, work in Britain and in other countries is given a good deal more notice than is sometimes the case in texts of this kind. This compilation of evidence makes the book a useful source of reference but it is inevitable in a book of this kind that many references cannot be given in much detail and some papers are only mentioned. In places, the book reads as a review of research. This raises the question of what audience it is intended for. At one level, the author assumes a considerable knowledge of theoretical and clinical psychology; at the other extreme, he gives brief explanations of reading readiness tests, projective techniques and the Vineland. The research worker and advanced student could use the book as a starting point for further reading, but the general student could have done with fuller accounts of selected research. The significance of work quoted is not always clear from the information given. Nevertheless, the book would have value in college libraries as a source of information and as a means of effecting an entry into the literature.

R. GULLIFORD.

SANDERS, C. (1962). *Psychological and Educational Bases of Academic Performance.* Australian Council for Educational Research, pp. xiv+165, 42s. Aust.

This volume reports the performance in school-leaving, and subsequent university examinations, of students entering the University of Western Australia in 1947, together with their scores for the ACER Test B42 (divided into verbal and non-verbal sub-tests), Raven's Matrices, and Mill-Hill Vocabulary. An almost identifical factor-structure emerged in separate arts and science analyses of the intercorrelations among the variables, with the school and university examinations, and the verbal test scores; clustering together closely to define a very strong factor, on which the non-verbal tests had negligible loadings. Another important analysis showed that the individual papers in both the arts and science first-year examinations must have high loadings on a general factor of 'educational aptitude.' The best single predictor to examination performances, in all faculties, proved to be the school-leaving average. Multiple regression equations, incorporating test data as well, produced only marginal improvements in predictive efficiency.

Numerous comparisons between sexes, faculties, and other sub-groups are presented, and one or two of the authors conclusions from these seem to be of dubious validity. For example, the intellectual level of 1947 entrants is compared with that of more recent students by using I.Q. equivalents based on two quite different tests. Again, the superior test scores and examination marks, of some arts students entering after serving in the Forces, are attributed to the influence of age, maturity, and experience, although the author himself points out elsewhere that these students were highly selected for educational aptitude, as measured by tests taken in the Forces.

The book is packed so full of useful data that it would inevitably be slightly indigestible, but poor editing makes the reader's task unnecessarily difficult. Many terms are defined almost every time they are used, the extreme example being on page 101 where the meaning of *civilian student* is explained three times in the space of twenty lines. Reliability sometimes means validity, and sometimes internal consistency, and asterisks sometimes denote results which are significant, and sometimes those which are not. The design of some of the Tables can be criticized, as for example Table 9, where faculty, intake and sex comparisons are all confounded in such a way that it is impossible to disentangle their several effects, so that it is necessary to search the text for the comparisons which are really required. Each cell in this table, moreover, contains an entry in brackets which is nowhere defined, although it is possible to deduce what it must be.

The concluding chapter provides a useful summary and synthesis, and in relating his findings to the results of other investigations, Professor Sanders develops a number of stimulating hypotheses which should provoke much further research.

The book presents the results of a valuable piece of research, and should certainly be read by all those who have a special interest in problems of selection at the university level.

W D. FURNEAUX.

SHIELDS, R. W. (1962). *A Cure of Delinquents*. Heinemann, pp. 188, 21s.

This book gives an account of the author's experience as a psycho-therapist at Bredinghurst School, while it was evolving as a unique educational and therapeutic environment. The book describes clearly both the problems of maladjusted children and the difficulties which beset the therapist in this setting.

The author summarises the guiding principles of his work in this way. He concludes that maladjustment and delinquency involve two basic elements, first actual deprivation in childhood and secondly, a current acute distortion in ego-development and personality integration. Any adequate treatment that aims at meeting and directing the disorders of maladjustment and delinquency must be designed to deal with these two basic elements. He suggests that treatment should include (*a*) making good the initial experiential deficiency, which is possible only in a stable setting—" it is here that educationists and psycho-therapists not only can but must work in close association "—and (*b*) the provision of adequate psychotherapy which alone can help the child to gain insight into and resolve his conflicts, unconscious anxieties, and reactive character traits.

Following on Winnicott's theoretical formulation of the problems, Dr. Shields shows how many individual cases respond when treated along analytic lines in an educational setting.

One of the most important contributions of the book is the clear way in which practical problems of administration and management are discussed. The conflicts between the educational, child care and therapy aspects of the work are clearly stated. Although Dr. Shields holds strong views on the importance of psychotherapy, he shows throughout the book an appreciation of the contributions of teachers and child care staff.

Among other aspects discussed are the relationships between those responsible for the structure of the community and those responsible for individual treatment, the relationship between maladjustment and delinquency, problems of discipline, of communication and of staff relationships. Even where one might disagree with the author as to the theoretical concepts underlying what he says, there is no doubt that the points he makes are practical and important.

This book should prove a useful addition to the studies available on the work of special schools for maladjusted children in this country. Its wider usefulness may be limited by the fact that it tends to fall between two stools ; informing the general reader about therapy in a special school setting and the discussion of analytic treatment for maladjusted children for those working in the field. For the former one feels that too much knowledge of analytic theory is assumed and for the latter the discussion may not be pushed far enough.

J. FISH.

WOODRUFF, A. D. (1961). *Basic Concepts of Teaching*. San Francisco: Chandler, pp. 238, $2.25.

The aim of this book is to present " a workable set of ideas of what a teacher does when he teaches a class." Having, in the introduction, outlined thirteen basic concepts, each with its supporting concepts, the author develops these systematically and consistently throughout the book, so that at the end, the student-teacher could feel with some justification that he had become acquainted with an intelligible and comprehensive set of principles, had received a good deal of specific guidance about how to translate them into practice, and, more important, had been encouraged and provoked into forming his own ideas of what teaching is all about.

The book, we are told, is part of a ' new program ' in teacher education at Utah State College of Education. It is meant as a first text-book, " before the solid courses in the substantive disciplines that constitute the foundations of education." The ' new program ' is very clearly a reaction against the excesses of so-called ' progressive education.' ' Activity methods,' ' problem solving approaches,' ' real life situations ' and ' education for mental health ' are out : in their place, " we can re-establish subject matter but on a conceptual not a verbal basis." The approach is in line with much that is generally accepted in this country to-day, although some readers may share the reviewer's feelings that parts of the book are written with the over-emphasis of the convert; for example, " that creative activity, critical thinking and problem solving are special activities which can be cultivated and used best at advanced levels of education," that "too much final determination of what is to be studied is left to the individual teacher," and that the activities of school societies " must be closely related to the content of the curriculum " to make them " educationally profitable."

Nevertheless, respect is due to the author for the skill of his detailed development of his main theme—that, " education is concerned with enabling people to make wise decisions " ; that values and feelings are therefore of primary importance but are themselves part and parcel of conceptual learning ; that the learning of symbols, skills and habits are important in so far as they help or hinder conceptual learning. This book is a cogent argument for the re-instatement of the intellect in those schools where it has been submerged in a program of activity for activity's sake. It is, incidentally, no less cogent a plea for the introduction of mere meaningful learning in those other schools where emphasis on examination results has resulted in a tendency to neglect any activity other than the cramming of factual information. E. FISK.

PUBLICATIONS RECEIVED

Mention of a book in this list neither implies, nor precludes, a later review.

J. A. Hadfield (1962): *Childhood and Adolescence.* Penguin Books, Ltd., pp. 286, 5s.

K. G. von Durckheim (1962). *Hara: The Vital Centre of Man.* George Allen and Unwin, Ltd., pp. 208, 21s.

C. Berg (1962). *Madkind—The Origin and Development of the Mind.* George Allen and Unwin, Ltd., pp. 277, 36s.

Anna Freud (Edit.), *et al.* (1962). *The Psycho-Analytic Study of the Child—Vol. XVI.* The Hogarth Press, pp. 563, 60s.

C. Sanders (1962). *Psychological and Educational Bases of Academic Performance.* Australian Council for Educational Research, pp. xiv+165, 42s. Aust.

Helen Huus (Edit.) (1961). *Education: Intellectual, Moral, Physical.* Oxford University Press, pp. 304, 30s.

J. W. Getzels and P. W. Jackson (1962). *Creativity and Intelligence.* John Wiley and Sons, pp. xvii+293, 49s.

L. S. Vygotsky (1962). *Thought and Language.* John Wiley and Sons, pp. xxi+168, 38s.

George Barnett and Jack Otis (1962). *Corporate Society and Education.* University of Michigan Press, pp. x+297, 50s.

Frank O. Copley (1962). *The American High School and the Talented Student.* University of Michigan Press, pp. xvii+92, 25s.

Barrington Kaye (1962). *Bringing up Children in Ghana.* George Allen and Unwin, pp. 244, 40s.

Norris G. Haring and E. Lakin Phillips (1962). *Educating Emotionally Disturbed Children.* McGraw-Hill, pp. xv+322, 50s. 6d.

Asahel D. Woodruff (1961). *Basic Concepts of Teaching.* Chandler Publishing Co., pp. ix+238, no price stated.

John G. McKenzie (1962). *Guilt—Its Meaning and Significance.* George Allen and Unwin, pp. viii+192, 21s.

Fillmore H. Sanford (1961). *Psychology: A Scientific Study of Man.* Wadsworth Publishing Co., California, pp. 569, no price stated.

Karl Mannheim and W. A. C. Stewart (1962). *An Introduction to the Sociology of Education.* Routledge and Kegan Paul, pp. xvii+187, 21s.

E. F. Carritt (1962). *The Theory of Beauty.* Methuen and Co., pp. vi+244, paperback edition, 9s. 6d.

M. D. Vernon (1962). *The Psychology of Perception.* Pelican Books, pp. 265, 5s.

Asahel D. Woodruff (1962). *Basic Concepts of Teaching.* Chandler Publishing Co., San Francisco, pp. x+300, no price stated.

Burton R. Clark (1962). *Educating the Expert Society.* Chandler Publishing Co., San Francisco, pp. xi+301, $2.75.

Jill Brandon and David Norris (1962). *Fountain Picture Books Nos.* 1, 2, 3. Ginn and Co., Ltd., pp. 32 each, no price stated.

Jerome S. Bruner (1962). *On Knowing—Essays for the Left Hand.* Harvard University Press; London: Oxford University Press, pp. 165, 30s.

George Z. F. Bereday and J. A. Lauwerys (Edits.) (1962). *The Year Book of Education,* 1962. Evans' Bros., Ltd., pp. xiv+541, 63s.

Frank A. Geldard (1962). *Fundamentals of Psychology.* John Wiley and Sons, pp. 437, 57s.

A. Taylor (Edit.) (1962). *Educational and Occupational Selection in West Africa.* Oxford University Press, pp. x+219, 25s.

Aubrey J. Yates (1962). *Frustration and Conflict.* Methuen and Co., Ltd., pp. x+236, 28s.

R. S. Peters (Edit.) (1962). *Brett's History of Psychology.* George Allen and Unwin, pp. 778, 60s.

J. A. Downing (1962). *Tw Bee or Not to Be.* Cassell and Co., Ltd., pp. xii+123, 9s. 6d.

Doris M. Lee (1962). *Background to Mathematical Development.* Oldbourne Press, pp. 227, 17s. 6d.

F. J. Fish (1962). *Schizophrenia.* John Wright and Sons, Ltd., pp. x+190, 32s. 6d.

Stuart Margulies and Lewis D. Eigen (1962). *Applied Programmed Instruction.* John Wiley and Sons, pp. ix+387, 53s.

Leopold Bellak (Edit.) (1962). *Contemporary European Psychiatry.* Grove Press Inc., pp. xxvi+372, $7.50.

F. H. George (1962). *Cognition.* Methuen and Co., Ltd., pp. 309, 32s. 6d.

Ira J. Gordon (1962). *Human Development.* Harper and Brothers, New York, pp. xiv+400, 46s.

CORRECTION

In the article "The Development of Formal Thinking in Verbal Comprehension," by D. Case and J. M. Collinson, the top line of page 110 should read: " Propositional judgments appear from these tests . . . "

Calcutta University Central Library

INDEX TO VOLUME XXXII, 1962

Articles :

Attitude Scale, A Further Approach to : R. S. Adams .. 201
Classification, Abstract and Concrete Modes of, in a Primitive Society : D. R. Price-Williams .. 50
Dyslexia, Specific : M. D. Vernon .. 143
Eleven Plus Classification, Examining Essays for : D. Young .. 267
Formal Thinking in Verbal Comprehension, The Development of : D. Case and J. M. Collinson .. 103
Gesell, Dr. Arnold, 1880-1961: L. B. Ames .. 101
Intelligence and Educational Attainment, The Relation of Neuroticism and Extraversion to : J. B. Biggs .. 188
Comments on the Article by J. B. Biggs : R. Lynn .. 196
Intelligence, The Effect of Town and Country upon Children's : H. B. Young, R. Tagiuri, G. Tesi and G. Montemagni .. 151
Learning a New Number System, A Factorial Study of, and its Relation to Attainment, Intelligence and Temperament : M. E. Hebron .. 38
Learning Theory, The Need for a Schematic : R. R. Skemp .. 133
Maladjusted Children, Residential Treatment of, A Study of some factors related to Progress in Adjustment : I. R. J. Petrie .. 29
Maudsley Personality Inventory, Three Experiments Using the Junior
  M. P. Callard and C. L. Goodfellow : I.—Neuroticism and Extraversion in Schoolboys as measured by the J M P.I. .. 241
  R. D. Savage : II.—Personality Factors and Academic Performance 251
  C. G. Costello and H. M. Brachman : III.—Cultural and Sex Differences in Extraversion and Neuroticism reflected in the responses to a Children's Personality Inventory .. 254
Motor Response, The Relationship between Inhibition of, and Cognitive Performance : C. C. Anderson .. 234
Musical Theme, Factors involved in Learning a Simple : G. E. R. Burroughs and J. N. Morris .. 18
Readability of Books of Low Difficulty, The Estimation of : J. McLeod .. 112
Reading Attainment and the Inconsistency of English Spelling at the Infant School Stage, The Relationship Between: J. A. Downing .. 166
Reading Skill at Eight years in Relation to Speech and other Factors : O. C. Sampson 12
Remedial Reading, A Pilot Experiment in, at The Hospital School, Lingfield, 1957-60 : L E. Keating .. 62
Remedial Teaching, A Summary of the Reading Ages of Children who had been given : K. Lovell, E. Johnson and O. Platts .. 66
School Phobia : M Chazan .. 209
Social Attitudes Among High School Students in Indonesia : B. W. Sumotirto .. 3
Socio-Economic Status, The Effect of, on the Scholastic Achievement of Sixth Grade Children : R L. Curry .. 46
Symposium : Educational Research To-day :
  E. A. Peel : I.—Priorities in Educational Research .. 178
  G. E. R. Burroughs : II —Co-operative Research in Institutes of Education 182
  S Wiseman : III.—The Tools of Research—Men and Machines 218
  W. D. Wall : IV.—Resistances, Organisation and Costs .. 224
Television Programmes, Intelligibility of Schools : A Royds .. 159
Teacher, Children's Evaluation of the Characteristics of the Good : P. H. Taylor .. 258
Teachers, Sources of Dissatisfaction among a group of : W. G. A. Rudd and S. Wiseman .. 275
Wechsler Intelligence Scale for Children, The, Applied to a Sample of London Primary School Children : S. Jones .. 119

Critical Note :

Comments on an Article by M. L. Kellmer Pringle : R. R. Dale .. 72
A Reply : M. L. Kellmer Pringle .. 73

Critical Notice :

Creativity and Intelligence by J. W. Getzels and P. W. Jackson : C. Burt .. 292

Research Note :

Note on the Problem of Inference in Educational Research : C. D. Hardie .. 75
A Reply : A. C. Aitken and D. N. Lawley .. 299

Summaries of Researches Reported in Degree Theses 78, 200, 300

Authors:

ADAMS, R. S.: A Further Approach to Attitude Scaling 201
AITKEN, A. C., and LAWLEY, D. N.: A Reply to: The Problem of Inference in Educational Research, by C. D. Hardie 299
AMES, L. B.: D. Arnold Gesell, 1880-1961 101
ANDERSON, C. C.: The Relationship between Inhibition of Motor Response and Cognitive Performance 234
BIGGS, J. B.: The Relation of Neuroticism and Extraversion to Intelligence and Educational Attainment 188
BURROUGHS, G. E. R.: Co-operative Research in Institutes of Education 182
BURROUGHS, G. E R. and MORRIS, J. N.: Factors Involved in Learning a Simple Musical Theme 18
BURT, C.: Critical Notice—Creativity and Intelligence, by J. W. Getzels and P. W. Jackson 292
CALLARD, M. P., and GOODFELLOW, C. L.: Neuroticism and Extraversion in Schoolboys as Measured by the Junior Maudsley Personality Inventory .. 241
CASE, D., and COLLINSON, J. M.: The Development of Formal Thinking in Verbal Comprehension 103
CHAZAN, M.. School Phobia 209
COSTELLO, C. G., and BRACHMAN, H. M.: Cultural and Sex Differences in Extraversion and Neuroticism reflected in responses to a Children's Personality Inventory 254
CURRY, R. L.: The Effect of Socio-Economic Status on the Scholastic Achievement of Sixth-Grade Children 46
DALE, R. R.: Critical Note—Comments on an Article by M. L. Kellmer Pringle .. 72
DOWNING, J. A.: The Relationship Between Reading Attainment and the Inconsistency of English Spelling at the Infants' School Stage 166
HARDIE, C. D.: Note on the Problem of Inference in Educational Research .. 75
HEBRON, M. E.: A Factorial Study of Learning a New Number System and its Relation to Attainment, Intelligence and Temperament 38
JONES, S.: The Wechsler Intelligence Scale for Children Applied to a Sample of London Primary School Children 119
KEATING, L E.: A Pilot Experiment in Remedial Reading at The Hospital School, Lingfield, 1957-60 62
KELLMER PRINGLE, M. L: Critical Note—A Reply to Comments by R R. Dale .. 73
LOVELL, K., JOHNSON, E., and PLATTS, O.: A Summary of a Study of the Reading Ages of Children who had been given Remedial Teaching 66
LYNN, R. Comments on the Article by J. B. Biggs 196
McLEOD, J.: The Estimation of Readability of Books of Low Difficulty 112
PEEL, E. A: Priorities in Educational Research 178
PETRIE, I R. J.: Residential Treatment of Maladjusted Children: A Study of Some Factors Related to Progress in Adjustment 29
PRICE-WILLIAMS, D. R.: Astract and Concrete Modes of Classification in a Primitive Society 50
ROYDS, A: Intdelligibility of Schools Television Programmes 159
RUDD, W. G. A. and WISEMAN, S.: Sources of Dissatisfaction among a Group of Teachers 275
SAMPSON, O. C.: Reading Skill at Eight Years in Relation to Speech and other Factors 12
SAVAGE, R. D.: Personality Factors and Academic Performance 251
SKEMP, R. R.: The Need for a Schematic Learning Theory 133
SUMOTIRTO, B. W. Social Attitudes Among High School Students in Indonesia .. 3
TAYLOR, P. H.: Children's Evaluations of the Characteristics of the Good Teacher .. 258
VERNON, M D.: Specific Dyslexia 143
WALL, W. D.: Resistances, Organisation and Costs 224
WISEMAN, S.: The Tools of Research—Men and Machines 218
YOUNG, D.: Examining Essays for Eleven Plus Classification 267
YOUNG, H. B., TAGIURI, R., TESI, G., and MONTEMAGNI, G.: Influence of Town and Country upon Children's Intelligence 151

NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained, or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length*. The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement*. Articles should be *preceded* by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need *not* be included, but successive sections of the article should receive numbered headings, e.g. :

I. INTRODUCTION.
II. OUTLINE OF PREVIOUS RESEARCH.
III. PLAN OF THE INVESTIGATION.
  (i) *Subjects and materials*.
  (ii) *Experimental Procedure*.
IV. STATEMENT OF RESULTS.
V. DISCUSSION AND CONCLUSIONS.
VI. REFERENCES.

4.—*Typing*. All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper. There should be margins oı at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs *, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.[4] When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—for any one author—by order of date. The following standard form of references should be observed :

DEARBORN, W. F., and ROTHNEY, J. W. M. (1941). *Predicting the Child's Development*. Cambridge, Mass. : Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)
BURT, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, 9, 55-71.

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full).

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table *and* a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text,

8 —*Corrections*. The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions (other than printer's errors).

9.—*Reprints*. Authors receive fifty copies of their papers free. Extra copies, in multiples of fifty, may be had at cost price if the order is given when the proof is returned.

OTHER PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology.
The British Journal of Psychology (Monograph Supplements).
The British Journal of Medical Psychology.
The British Journal of Statistical Psychology
The British Journal of Social and Clinical Psychology.

The British Journal of Psychology and *The British Journal of Medical Psychology* appear in quarterly parts in volumes of about 350 pages, subscription price of each, payable in advance, 60s. per volume (post free), Subscriptions may be sent to any Bookseller or to the Cambridge University Press, Bentley House, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $10.00. Enquiries should be addressed to the Cambridge University Press, American Branch, 32, East 57th Street, New York 22.

The British Journal of Statistical Psychology is issued twice a year, in May and November, in volumes of about 184 pages. The subscription rate payable in advance is 30s. 6d. per volume (post free). The subscription price in the U.S.A. is $4.50 (post free). Subscriptions should be sent to the Society's Agents, Taylor and Francis, Ltd., Red Lion Court, Fleet Street, E.C.4., or to any Bookseller.

The British Journal of Social and Clinical Psychology is issued three times a year in February, June and November, in volumes of approximately 240 pages. The subscription price of each, payable in advance, is 50s. (post free). Subscriptions may be sent to any Bookseller, or to the Cambridge University Press, 200, Euston Road, London, N.W.1. The subscription price in the U.S.A. is $8.50. Enquiries should be sent to the Cambridge University Press American Branch, 32, East 57th Street, New York 22. Separate parts will be sold at 20s.: $3.50 in the United States.

A series of *Monograph Supplements* is issued in connection with the *British Journal of Psychology*, but is not included in the subscription for that journal. Orders may be sent to any Bookseller or to the Cambridge University Press, 200, Euston Road, London, N.W.1 from whom details of these publications may be obtained.

MANUSCRIPTS

For the *British Journal of Psychology* should be sent to:
Dr. Boris Semeonoff, Department of Psychology, The University, Edinburgh.

For the *British Journal of Medical Psychology*, should be sent to Dr Thomas Freeman, Lansdowne Clinic, 4, Royal Crescent, Kelvingrove Glasgow, C 3

For the *British Journal of Educational Psychology* should be sent to Mr. L. B. Birch, Institute of Education, The University, Sheffield, 10.

For the *British Journal of Statistical Psychology*, should be sent to the Assistant Editor, Sir Cyril Burt, 9, Elsworthy Road, Hampstead, London, N.W.3

For the *British Journal of Social and Clinical Psychology* should be sent to Mr. Michael Argyle (Social Psychology Editor), Institute of Experimental Psychology, 1, South Parks Road, Oxford;

or to:

Dr. Jack Tizard (Clinical Psychology Editor), Maudsley Hospital, Denmark Hill, London S.E.5.

For the *Monograph Supplements*, to Professor L. S. Hearnshaw, Department of Psychology, The University, 7, Abercromby Square, Liverpool, 7, whose decision regarding the publication of all *Monograph Supplements* will be final. No MS. can be considered for publication in any year unless it is in the hands of the Editor by 1st September of the preceding year.

OFF-PRINTS

Contributors receive 50 free copies of their article and may buy more in units of 50.

Members of the British Psychological Society receive all the Society's Journals at reduced rates.

Information on membership and further activities may be obtained from the Secretary, Tavistock House South, Tavistock Square, London, W.C.1.

PUBLISHED BY THE ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION

Education for Teaching is published three times a year at mid-term in November, February and May and distributed free to members of the Association. To non-members it costs, 5s. per issue, post free (subscription 14s. per annum).

Members of the A.T.C.D.E. receive the *British Journal of Educational Psychology* at a reduced rate.

Information on membership and further activities may be obtained from the Honorary Secretary, 151, Gower Street, London, W.C.1.

2 JAN 1963

CONTENTS

PAGE

R. S. ADAMS : A Further Approach to Attitude Scaling . . . 201

M. CHAZAN : School Phobia 209

SYMPOSIUM : EDUCATIONAL RESEARCH TODAY :

S. WISEMAN : III.—The Tools of Research—Men and Machines . 218

W. D. WALL : IV.—Resistances, Organisation and Costs . . 224

C. C. ANDERSON : The Relationship between Inhibition of Motor Response and Cognitive Performance 234

THREE EXPERIMENTS USING THE JUNIOR MAUDSLEY PERSONALITY INVENTORY :

M. P. CALLARD and C. L. GOODFELLOW : I.—Neuroticism and Extraversion in Schoolboys as Measured by the J.M.P.I. 241

R. D. SAVAGE : II.—Personality Factors and Academic Performance. 251

C. G. COSTELLO and H. M. BRACHMAN : III.—Cultural and Sex Differences in Extraversion and Neuroticism reflected in responses to a Children's Personality Inventory 254

P. H. TAYLOR : Children's Evaluations of the Characteristics of the Good Teacher 258

D. YOUNG : Examining Essays for Eleven Plus Classification . 267

W. G. A. RUDD and S. WISEMAN : Sources of Dissatisfaction among a Group of Teachers. 275

CRITICAL NOTICE :

C. BURT : Creativity and Intelligence, by J. W. Getzels and P. W. Jackson 292

RESEARCH NOTE :

A. C. AITKEN and D. N. LAWLEY : A reply to : The Problem of Inference in Educational Research, by C. D. Hardie 299

SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES 300

BOOK REVIEWS 311

PUBLICATIONS RECEIVED 322

INDEX TO VOLUME XXXII 324

NOTES TO CONTRIBUTORS 326